

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 15:05:25 ; Search time 39.0811 Seconds
(without alignments)
4076.547 Million cell updates/sec

Title: US-10-665-283-1

Perfect score: 9903
Sequence: 1 MALRGCASAGSDPLMDMNV.....RSVAVAKKPKFSISPSLS 1927

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5 COMB.pep: *
2: /cgn2_6/ptodata/1/iaa/6 COMB.pep: *
3: /cgn2_6/ptodata/1/iaa/H COMB.pep: *
4: /cgn2_6/ptodata/1/iaa/PTUS COMB.pep: *
5: /cgn2_6/ptodata/1/iaa/RE COMB.pep: *
6: /cgn2_6/ptodata/1/iaa/backfillseq.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7860	79.4	1531	1 US-08-463-092B-4	Sequence 4, Appli
2	7860	79.4	1531	1 US-08-463-109A-4	Sequence 4, Appli
3	7860	79.4	1531	1 US-08-460-907B-4	Sequence 4, Appli
4	7860	79.4	1531	2 US-08-463-179A-4	Sequence 4, Appli
5	7860	79.4	1531	2 US-08-461-384B-4	Sequence 4, Appli
6	7860	79.4	1531	2 US-09-647-140B-19	Sequence 19, Appli
7	7849	79.3	1531	1 US-08-141-893-2	Sequence 2, Appli
8	7849	79.3	1531	1 US-08-463-092B-2	Sequence 2, Appli
9	7849	79.3	1531	1 US-08-463-109A-2	Sequence 2, Appli
10	7849	79.3	1531	1 US-08-460-907B-2	Sequence 2, Appli
11	7849	79.3	1531	2 US-08-463-179A-2	Sequence 2, Appli
12	7849	79.3	1531	2 US-08-461-384B-2	Sequence 2, Appli
13	7849	79.3	1531	2 US-08-407-207A-2	Sequence 2, Appli
14	7002.5	70.7	1528	1 US-08-463-092B-6	Sequence 6, Appli
15	7002.5	70.7	1528	1 US-08-463-109A-6	Sequence 6, Appli
16	7002.5	70.7	1528	1 US-08-460-907B-6	Sequence 6, Appli
17	7002.5	70.7	1528	2 US-08-463-179A-6	Sequence 6, Appli
18	7002.5	70.7	1528	2 US-08-461-384B-6	Sequence 6, Appli
19	4487.5	45.3	1527	2 US-09-647-140B-6	Sequence 6, Appli
20	4487.5	45.3	1527	2 US-09-647-140B-33	Sequence 33, Appli
21	3391.5	34.2	1503	2 US-09-647-140B-8	Sequence 8, Appli
22	3385.5	34.2	1503	2 US-09-792-616-3	Sequence 3, Appli
23	3218.5	32.5	1498	2 US-09-792-616-9	Sequence 9, Appli
24	2374.5	24.0	1621	2 US-08-972-927-3	Sequence 3, Appli
25	2349.5	23.7	1622	2 US-08-972-927-6	Sequence 6, Appli
26	2266.5	22.9	1325	2 US-09-647-140B-2	Sequence 2, Appli
27	2212.5	22.3	1464	2 US-10-012-896-1008	Sequence 1008, Ap

28	2195.5	22.2	1261	2 US-09-439-313-538	Sequence 538, App
29	2195.5	22.2	1261	2 US-09-636-215-538	Sequence 538, App
30	2195.5	22.2	1261	2 US-09-685-166A-538	Sequence 538, App
31	2195.5	22.2	1261	2 US-09-679-426-538	Sequence 538, App
32	2195.5	22.2	1261	2 US-09-759-143-538	Sequence 538, App
33	2195.5	22.2	1261	2 US-09-651-236-538	Sequence 538, App
34	2195.5	22.2	1261	2 US-09-657-279-538	Sequence 538, App
35	2195.5	22.2	1261	2 US-10-012-896-538	Sequence 538, App
36	2195.5	22.2	1261	2 US-10-012-896-1009	Sequence 1009, App
37	2118	21.4	1581	2 US-08-726-320-3	Sequence 3, Appli
38	2118	21.4	1581	2 US-09-208-716-3	Sequence 3, Appli
39	2115	21.4	1228	2 US-09-439-313-537	Sequence 537, App
40	2115	21.4	1228	2 US-09-636-215-537	Sequence 537, App
41	2115	21.4	1228	2 US-09-685-166A-537	Sequence 537, App
42	2115	21.4	1228	2 US-09-679-426-537	Sequence 537, App
43	2115	21.4	1228	2 US-09-759-143-537	Sequence 537, App
44	2115	21.4	1228	2 US-09-651-236-537	Sequence 537, App
45	2115	21.4	1228	2 US-09-657-279-537	Sequence 537, App

ALIGNMENTS

RESULT 1
US-08-463-092B-4
Sequence 4, Application US/08463092B
Patent No. 576680
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
MULTIDRUG RESISTANCE PROTEINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,092B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: Q1546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1531 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-463-092B-4

Query Match 79.4%; Score 7860; DB 1; Length 1531;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MALRGFCADGSDPLMDNNVTWNTSNPDPTKCFONTVLVWPCFYLMACFPFFLYLSRH 60
DB 1 MALRGFCADGSDPLMDNNVTWNTSNPDPTKCFONTVLVWPCFYLMACFPFFLYLSRH 60
QY 61 DRGYIOMPLUNKTKTALGFILMIYCMADLFYSFMRSRSGIFLAPVFLVSPILLGITLLA 120
DB 61 DRGYIOMPLUNKTKTALGFILMIYCMADLFYSFMRSRSGIFLAPVFLVSPILLGITLLA 120
QY 61 DRGYIOMPLUNKTKTALGFILMIYCMADLFYSFMRSRSGIFLAPVFLVSPILLGITLLA 120
DB 61 DRGYIOMPLUNKTKTALGFILMIYCMADLFYSFMRSRSGIFLAPVFLVSPILLGITLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFLVALVQALALRSKIMTALKEDAQVDLFRDITFVYVS 180
DB 121 TFLIQLERRKGVSSGIMLTFLVALVQALALRSKIMTALKEDAQVDLFRDITFVYVS 180
QY 121 TFLIQLERRKGVSSGIMLTFLVALVQALALRSKIMTALKEDAQVDLFRDITFVYVS 180
DB 121 TFLIQLERRKGVSSGIMLTFLVALVQALALRSKIMTALKEDAQVDLFRDITFVYVS 180
QY 181 LLLIQLVISCPSDNRSLFSETHIDNPPCPSSASFLSRITFWMITGLIVRGYRPLEGSD 240
DB 181 LLLIQLVISCPSDNRSLFSETHIDNPPCPSSASFLSRITFWMITGLIVRGYRPLEGSD 240
QY 241 LMSLKEDTSEOVVAVLVKNNKKECAKTRKOPVKVYVSSKDPAPCKESSKYDANEVVAL 300
DB 241 LMSLKEDTSEOVVAVLVKNNKKECAKTRKOPVKVYVSSKDPAPCKESSKYDANEVVAL 300
QY 301 IVKSPQKEMNPSLFVLYKTGPFYLMGFFFKAIHDLMMFGSPQILKLIKRVNDTKAPD 360
DB 301 IVKSPQKEMNPSLFVLYKTGPFYLMGFFFKAIHDLMMFGSPQILKLIKRVNDTKAPD 360
QY 361 WQGYFYVLLFYTAQLQTLVHQYHICFVSGMRKIKTAVIGAVYKALVITNSARKSTV 420
DB 361 WQGYFYVLLFYTAQLQTLVHQYHICFVSGMRKIKTAVIGAVYKALVITNSARKSTV 420
QY 421 GEIVNLMVDARFMDLATYINMISAPLOVTLAYLLMLNIGSPVLAVGVAMVLMVNVN 480
DB 421 GEIVNLMVDARFMDLATYINMISAPLOVTLAYLLMLNIGSPVLAVGVAMVLMVNVN 480
QY 481 AVMAKTKTYOVAHMKSKDNRIKLNNEILNGIKVLKLYAMELAFDKYLAIRQEBLKYLK 540
DB 481 AVMAKTKTYOVAHMKSKDNRIKLNNEILNGIKVLKLYAMELAFDKYLAIRQEBLKYLK 540
QY 541 KSAVISAAGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLAFNIIKRPPLNII 600
DB 541 KSAVISAAGTFTWCTPFLVALCTFAVYVTTIDENNILDAQTAFAVSLAFNIIKRPPLNII 600
QY 601 MVISSIVASVSLKRLIFLSHEELPDSIERPRKDGCTNSITVRNATFTWASDDPT 660
DB 601 MVISSIVASVSLKRLIFLSHEELPDSIERPRKDGCTNSITVRNATFTWASDDPT 660
QY 661 LINGIFFSIPEGALVAVVGVCGCKSSLSALLAEMDKYGHVAIKGSVAVYVQAAMIOND 720
DB 661 LINGIFFSIPEGALVAVVGVCGCKSSLSALLAEMDKYGHVAIKGSVAVYVQAAMIOND 720
QY 721 SLRENILFCQLEBPYRSVIOACALLPDLEILPSGDRTEIGEKVNLSSGQKQVSLAR 780
DB 721 SLRENILFCQLEBPYRSVIOACALLPDLEILPSGDRTEIGEKVNLSSGQKQVSLAR 780
QY 781 AVYSNADTYLPDDPLSAVDANVGHKIFENVIGPKMLKNKRIIIVTHSSSYLPQVDVIV 840
DB 781 AVYSNADTYLPDDPLSAVDANVGHKIFENVIGPKMLKNKRIIIVTHSSSYLPQVDVIV 840
QY 841 MSGGKISEMGSYOELLARDGAPAFELRTASTEOBDAEENGVTGSGPGKEAKQEMNGM 900
DB 841 MSGGKISEMGSYOELLARDGAPAFELRTASTEOBDAEENGVTGSGPGKEAKQEMNGM 900
QY 901 LVTSAGKOLQROUSSSSSYSGDISRHNSTAELQKAEKKETWLMEDAKQOTGOYKL 960
DB 901 LVTSAGKOLQROUSSSSSYSGDISRHNSTAELQKAEKKETWLMEDAKQOTGOYKL 960

```

```

DB 901 LVTSAGKOLQROUSSSSSYSGDISRHNSTAELQKAEKKETWLMEDAKQOTGOYKL 960
QY 961 SVYWDYMKAIQGLFISFISFLFMCNHYSAASNYWLSLMTDDPIVNTQOHTKYRLSVYG 1020
DB 961 SVYWDYMKAIQGLFISFISFLFMCNHYSAASNYWLSLMTDDPIVNTQOHTKYRLSVYG 1020
QY 1021 ALGISGCIAVFGYMAVSIIGIILASRCLHYDLHSLIRSMSFFERTPSGNLVNRSEKEL 1080
DB 1021 ALGISGCIAVFGYMAVSIIGIILASRCLHYDLHSLIRSMSFFERTPSGNLVNRSEKEL 1080
QY 1081 DTYSMTPEVITKMGSLFNVIGACIYIILATPIAIIIPPLGIYFVFQRPVYASSROL 1140
DB 1081 DTYSMTPEVITKMGSLFNVIGACIYIILATPIAIIIPPLGIYFVFQRPVYASSROL 1140
QY 1141 KRLSVSRSPVYSHFNTELLGVSYIRAFEOERFIHOSDLKVDENQAYYPSIVANWELA 1200
DB 1141 KRLSVSRSPVYSHFNTELLGVSYIRAFEOERFIHOSDLKVDENQAYYPSIVANWELA 1200
QY 1201 VRLECVGNCIYLFALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETNIVA 1260
DB 1201 VRLECVGNCIYLFALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMIOETAPSSSWPOYGRVFRVYCLRYREDLDPVLRIHVVITNGG 1320
DB 1261 VERLKEYSETEKAPMIOETAPSSSWPOYGRVFRVYCLRYREDLDPVLRIHVVITNGG 1320
QY 1321 EKVGIVRTGAGKSLTLGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIVRTGAGKSLTLGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGSLRMNLDPEFSQYSDIEWWTSLELAHLKOPVSAIPLKLDHECAGEGENTSVGQROLVCL 1440
DB 1381 SGSLRMNLDPEFSQYSDIEWWTSLELAHLKOPVSAIPLKLDHECAGEGENTSVGQROLVCL 1440
QY 1441 ARALLRRTKILVDEAPRAVDLETFDDLIQSTIRFOEDCTVLTIAHRLNTIMDTRYIVL 1500
DB 1441 ARALLRRTKILVDEAPRAVDLETFDDLIQSTIRFOEDCTVLTIAHRLNTIMDTRYIVL 1500
QY 1501 DKGEIOEYGAPSDILQOQGLFYSMAKXAGLV 1531
DB 1501 DKGEIOEYGAPSDILQOQGLFYSMAKXAGLV 1531

```

RESULT 2
 US-08-462-109A-4
 ; Sequence 4, Application US/08462109A
 ; Patent No. 5882875
 ; GENERAL INFORMATION:
 ; APPLICANT: Cole, Susan P.C.
 ; APPLICANT: Deeley, Roger G.
 ; TITLE OF INVENTION: METHODS FOR IDENTIFYING
 ; TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD
 ; STREET: 60 State Street, suite 510
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/462,109A
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/966,923
 ; FILING DATE: 27-OCT-1992
 ; APPLICATION NUMBER: 08/029,340

FILING DATE: 8-MAR-1993
 APPLICATION NUMBER: 08/141,893
 FILING DATE: 26-OCT-1993
 APPLICATION NUMBER: 08/407,207
 FILING DATE: 20-MAR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: DeConti, Giulio A. Jr.
 REGISTRATION NUMBER: 31,503
 REFERENCE/DOCKET NUMBER: P01-002CP4
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1531 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-462-109A-4

Query Match 79.4%; Score 7860; DB 1; Length 1531;
 Best Local Similarity 100.0%; Pred. NO. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNVNTMTSNPDFTKCFONTVLVWPCFYLMACFPFYLYSRH 60
 DB 1 MALRGFCSADGSDPLMDMNVNTMTSNPDFTKCFONTVLVWPCFYLMACFPFYLYSRH 60
 QY 61 DRGTYQMTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLADVFLVPTLLGITTLLA 120
 DB 61 DRGTYQMTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLADVFLVPTLLGITTLLA 120
 QY 121 TPLIQERRRGVSSGIMLTFMVALYCALAIIASKIMTALKEAOVDLPDITFYYPFS 180
 DB 121 TPLIQERRRGVSSGIMLTFMVALYCALAIIASKIMTALKEAOVDLPDITFYYPFS 180
 QY 121 TPLIQERRRGVSSGIMLTFMVALYCALAIIASKIMTALKEAOVDLPDITFYYPFS 180
 DB 121 TPLIQERRRGVSSGIMLTFMVALYCALAIIASKIMTALKEAOVDLPDITFYYPFS 180
 QY 181 LMLQLVLSGFSRSPFSETIHDNPNCPRESSASFLSRITFWMTGILVRYRPLGSD 240
 DB 181 LMLQLVLSGFSRSPFSETIHDNPNCPRESSASFLSRITFWMTGILVRYRPLGSD 240
 QY 241 LMSLNKEDTSEQVVPVLVKNWKKCAKTRKQPVVVYSSKDPAPKSSKVDANEVEAL 300
 DB 241 LMSLNKEDTSEQVVPVLVKNWKKCAKTRKQPVVVYSSKDPAPKSSKVDANEVEAL 300
 QY 241 LMSLNKEDTSEQVVPVLVKNWKKCAKTRKQPVVVYSSKDPAPKSSKVDANEVEAL 300
 DB 241 LMSLNKEDTSEQVVPVLVKNWKKCAKTRKQPVVVYSSKDPAPKSSKVDANEVEAL 300
 QY 301 IVSPQKKNPSPLEKVLKTFGPFLMSFPFKAIHDMFSGPQIILKIKFVDTKAPD 360
 DB 301 IVSPQKKNPSPLEKVLKTFGPFLMSFPFKAIHDMFSGPQIILKIKFVDTKAPD 360
 QY 361 MOGFYFVLLFVTAQCLQTLVHOYFHI CFVSGMR IKTAVIGAVYRKALVITNSARKSTV 420
 DB 361 MOGFYFVLLFVTAQCLQTLVHOYFHI CFVSGMR IKTAVIGAVYRKALVITNSARKSTV 420
 QY 421 GEIVNLMSVDAQRMDLATYINMISAPLOVYLLALYMLNGPSVLAVGAVVWLVMPVN 480
 DB 421 GEIVNLMSVDAQRMDLATYINMISAPLOVYLLALYMLNGPSVLAVGAVVWLVMPVN 480
 QY 481 AYWAMKTKTYQVAMKSKDNRIKIMNEILINGIKLYAMELAFKDYLAIROBELVYLK 540
 DB 481 AYWAMKTKTYQVAMKSKDNRIKIMNEILINGIKLYAMELAFKDYLAIROBELVYLK 540
 QY 541 KSAVLSAVGTFTWCTCFELVALCTFAYVYTIIDENNIIDAOAFSLAFNILRPLNLP 600
 DB 541 KSAVLSAVGTFTWCTCFELVALCTFAYVYTIIDENNIIDAOAFSLAFNILRPLNLP 600
 QY 601 MVISSIVQASVSLKRLRIFLSHEBELDPSIEREPVKDGGGINSITVNNATFTWARSPT 660
 DB 601 MVISSIVQASVSLKRLRIFLSHEBELDPSIEREPVKDGGGINSITVNNATFTWARSPT 660
 QY 661 LNTGTFSPREGALVAVVGVGCGSSLSLALLAEMDVKEGVHAIKGSVAVYPOQAWIQND 720
 DB 661 LNTGTFSPREGALVAVVGVGCGSSLSLALLAEMDVKEGVHAIKGSVAVYPOQAWIQND 720
 QY 721 SLRENILFGQLEPEYRSVYIQCALLPDEILIPSGRTEIGEGVNLSSGQKQVSLAR 780

DB 721 SLRENILFGQLEPEYRSVYIQCALLPDEILIPSGRTEIGEGVNLSSGQKQVSLAR 780
 QY 781 AVYSNADITYFPDPLSADVDAVGHIFENYIGPKMKLXKTRILLVTHSMGYLPQVDYIIV 840
 DB 781 AVYSNADITYFPDPLSADVDAVGHIFENYIGPKMKLXKTRILLVTHSMGYLPQVDYIIV 840
 QY 841 MSGKISEMSYOELLARDGAPAEPLRTVASTQEOAEENGTVGSGPKKAEKQNGM 900
 DB 841 MSGKISEMSYOELLARDGAPAEPLRTVASTQEOAEENGTVGSGPKKAEKQNGM 900
 QY 901 LVYDSACKQLOQROSSSSYSGDISRRHNSYAELOKAEKKEETKLMLEADKQOTGVKL 960
 DB 901 LVYDSACKQLOQROSSSSYSGDISRRHNSYAELOKAEKKEETKLMLEADKQOTGVKL 960
 QY 961 SVYWDYKAIQGLFISPLSIFLPMCNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
 DB 961 SVYWDYKAIQGLFISPLSIFLPMCNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
 QY 1021 ALGISOGIAVGYSMASIGIILASRCLVHDLHSIRSPSPFERTPSGNLVNREKEL 1080
 DB 1021 ALGISOGIAVGYSMASIGIILASRCLVHDLHSIRSPSPFERTPSGNLVNREKEL 1080
 QY 1081 DTVDSMIPREVIKMFMSLFNVIAGCTVILATPIAIIIPPLGIYFFVORFYVASSROL 1140
 DB 1081 DTVDSMIPREVIKMFMSLFNVIAGCTVILATPIAIIIPPLGIYFFVORFYVASSROL 1140
 QY 1141 KRLSVRSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQAYVPSIVANRWLA 1200
 DB 1141 KRLSVRSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQAYVPSIVANRWLA 1200
 QY 1201 VRLCVCNCIVLFAALFAVIRSRSLAGVLSVSYSLQVTTVLMVLNRSSEMETNIVA 1260
 DB 1201 VRLCVCNCIVLFAALFAVIRSRSLAGVLSVSYSLQVTTVLMVLNRSSEMETNIVA 1260
 QY 1261 VERLKESETEKAPMOIOETAPSSWPQVGRVEFRNYCARYBEDDFVLRHINVTJNGG 1320
 DB 1261 VERLKESETEKAPMOIOETAPSSWPQVGRVEFRNYCARYBEDDFVLRHINVTJNGG 1320
 QY 1321 EKVIGVGTAGKSSLTGLFRINESABEIIIDGINIAKIGLHDLRFKTIIPDPVLF 1380
 DB 1321 EKVIGVGTAGKSSLTGLFRINESABEIIIDGINIAKIGLHDLRFKTIIPDPVLF 1380
 QY 1381 SGLSRNMLDPFSQYSDBEVWTSLELAHKDFVSLPDKLDHBCAEGGENLSVGQROLVCL 1440
 DB 1381 SGLSRNMLDPFSQYSDBEVWTSLELAHKDFVSLPDKLDHBCAEGGENLSVGQROLVCL 1440
 QY 1441 ARALLKRTKTLVDEATPAVDLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRYIVL 1500
 DB 1441 ARALLKRTKTLVDEATPAVDLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRYIVL 1500
 QY 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531
 DB 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531

RESULT 3
 US-08-460-907B-4
 : Sequence 4, Application US/08460907B
 : Patent No. 5891724
 : GENERAL INFORMATION:
 : APPLICANT: Deeley, Roger G.
 : APPLICANT: Cole, Susan P.C.
 : TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG
 : RESISTANCE ON A CELL
 : NUMBER OF SEQUENCES: 9
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: PARTIQ RESEARCH & DEVELOPMENT INNOVATIONS
 : STREET: Queen's University at Kingston
 : CITY: Kingston
 : STATE: Ontario
 : COUNTRY: CANADA
 : ZIP: K7L 3N6

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,907B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01551
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-907B-4

Query Match      79.4%; Score 7860; DB 1; Length 1531;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1 MALRGFCAGDSDPLMDNMTNTNSPDPFTKCFQNTLVWVPCFYLMACFPFYFLYSRH 60
Db      1 MALRGFCAGDSDPLMDNMTNTNSPDPFTKCFQNTLVWVPCFYLMACFPFYFLYSRH 60
Oy      61 DRGYIOMTPLNKTKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPTLGITLLA 120
Db      61 DRGYIOMTPLNKTKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPTLGITLLA 120
Oy      121 TFLIQLERRKGVQSSGIMLTFWLVALVCAALILRSKITALKEDQVDFRDIITYYFES 180
Db      121 TFLIQLERRKGVQSSGIMLTFWLVALVCAALILRSKITALKEDQVDFRDIITYYFES 180
Oy      121 TFLIQLERRKGVQSSGIMLTFWLVALVCAALILRSKITALKEDQVDFRDIITYYFES 180
Db      121 TFLIQLERRKGVQSSGIMLTFWLVALVCAALILRSKITALKEDQVDFRDIITYYFES 180
Oy      181 LLLIQLVSCPSDPSPLFSETIHDNPPCESASPLSRTTFWMLTGLIVRGYROPLEGSD 240
Db      181 LLLIQLVSCPSDPSPLFSETIHDNPPCESASPLSRTTFWMLTGLIVRGYROPLEGSD 240
Oy      241 LMSLNKEDTSQVAVPVLVKMKKECAKTRKQPVKVYSSKDPAPCKESSKYDANEVEAL 300
Db      241 LMSLNKEDTSQVAVPVLVKMKKECAKTRKQPVKVYSSKDPAPCKESSKYDANEVEAL 300
Oy      301 IVKSPQKMNNSLFLVLYKTGPFYLMGFFFAKIHDLMMFGSPQILKLIFVNDTKAPD 360
Db      301 IVKSPQKMNNSLFLVLYKTGPFYLMGFFFAKIHDLMMFGSPQILKLIFVNDTKAPD 360
Oy      361 MGGFYTYLTFYTAQLOTLVLYHQVFIQFVSGMRITKTAIVIGAVYKALVITNSAKKSTV 420
Db      361 MGGFYTYLTFYTAQLOTLVLYHQVFIQFVSGMRITKTAIVIGAVYKALVITNSAKKSTV 420
Oy      421 GEIVNLMSVDAQRFMDLATYINMIWSAPLQVILALYLLMLNLGSPVLGAVAVMVLMEVN 480
```

```

Db      421 GEIVNLMSVDAQRFMDLATYINMIWSAPLQVILALYLLMLNLGSPVLGAVAVMVLMEVN 480
Oy      481 AVAMAKTKTYQVAHMSKDKRIRIKLMEIINGIKVLKLYAMELAFKDKVLAIRBELKVLK 540
Db      481 AVAMAKTKTYQVAHMSKDKRIRIKLMEIINGIKVLKLYAMELAFKDKVLAIRBELKVLK 540
Oy      541 KSAVLSAVGFTFWCTPFLVACTFAVYVTIDENNIIDAQTAFAVSLAFNILRFPNLILP 600
Db      541 KSAVLSAVGFTFWCTPFLVACTFAVYVTIDENNIIDAQTAFAVSLAFNILRFPNLILP 600
Oy      601 MVISIVQASVSLKRLRIFLSHELEPDSIERRPVQGGGTNSITVNAFTWASDPPT 660
Db      601 MVISIVQASVSLKRLRIFLSHELEPDSIERRPVQGGGTNSITVNAFTWASDPPT 660
Oy      661 LMGITFSIPGALVAVAVGVQCGKSSLSLALMBMDKVEGHVAIKGSVAVPOQAMTQND 720
Db      661 LMGITFSIPGALVAVAVGVQCGKSSLSLALMBMDKVEGHVAIKGSVAVPOQAMTQND 720
Oy      721 SLRENILFGCQLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSCGQKQSVLSAR 780
Db      721 SLRENILFGCQLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSCGQKQSVLSAR 780
Oy      781 AVYSNADIYLPDDPLSAVDAAHVGKHFENYIGPGMKLKNKTRILYTHSMGYLPQVDVYIV 840
Db      781 AVYSNADIYLPDDPLSAVDAAHVGKHFENYIGPGMKLKNKTRILYTHSMGYLPQVDVYIV 840
Oy      841 MSGGKISEMSGYOELLARDGAFAEFLRTVASTEOBDAEENGVTGVSQPKCAKOMENGM 900
Db      841 MSGGKISEMSGYOELLARDGAFAEFLRTVASTEOBDAEENGVTGVSQPKCAKOMENGM 900
Oy      901 LVTDASAKOLQROLSSSSSYSGDISRRHNSYAEIOKAKEAKETWKLMEADKAQTGVKL 960
Db      901 LVTDASAKOLQROLSSSSSYSGDISRRHNSYAEIOKAKEAKETWKLMEADKAQTGVKL 960
Oy      961 SVYWDYKKAIGLFISELSIFLPMCNHVSALASNYLSLMTDDPVPNGTOHTKRLSVYG 1020
Db      961 SVYWDYKKAIGLFISELSIFLPMCNHVSALASNYLSLMTDDPVPNGTOHTKRLSVYG 1020
Oy      1021 ALGISOGIAVFGVMAVSIIGIILASRCLHYDLHSILRSPSPFERTPSGNLVNRFSEKL 1080
Db      1021 ALGISOGIAVFGVMAVSIIGIILASRCLHYDLHSILRSPSPFERTPSGNLVNRFSEKL 1080
Oy      1081 DTVDMSIPEVILKMGMSLFPNVIGACIVILLATPIAIIIPPLGIYFVQRFVYASSROL 1140
Db      1081 DTVDMSIPEVILKMGMSLFPNVIGACIVILLATPIAIIIPPLGIYFVQRFVYASSROL 1140
Oy      1141 KRLESVSRSPYSHFNFTLLGVSYTRAFEOERTIHOSDLKVDENOKAYPSIYANRWLA 1200
Db      1141 KRLESVSRSPYSHFNFTLLGVSYTRAFEOERTIHOSDLKVDENOKAYPSIYANRWLA 1200
Oy      1201 VRLECVNGCIYLPALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETNIVA 1260
Db      1201 VRLECVNGCIYLPALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETNIVA 1260
Oy      1261 VERLKEYSETEKEAPMOIOTETAPSSWPQVGRVFRNYCLRYREDLDFVLRHINVTINGG 1320
Db      1261 VERLKEYSETEKEAPMOIOTETAPSSWPQVGRVFRNYCLRYREDLDFVLRHINVTINGG 1320
Oy      1321 EKVGIVGRTGAKSSTLGLFRINESABGEIINDGINIAKIGLHDLRFKTIIIQDPVLF 1380
Db      1321 EKVGIVGRTGAKSSTLGLFRINESABGEIINDGINIAKIGLHDLRFKTIIIQDPVLF 1380
Oy      1381 SGRIRMLNDPSQVSDDEWMTSLELAHLKQFVSLPKLTHECAEGENLISVGOROLVCL 1440
Db      1381 SGRIRMLNDPSQVSDDEWMTSLELAHLKQFVSLPKLTHECAEGENLISVGOROLVCL 1440
Oy      1441 ARALIRKTKILIVDEATAAVDFETDDIIGSTIRQFEDCVLTIAHNLNTIMDTRVIVL 1500
Db      1441 ARALIRKTKILIVDEATAAVDFETDDIIGSTIRQFEDCVLTIAHNLNTIMDTRVIVL 1500
Oy      1501 DKGEIQEYGAESDILQORGLFYMAKDAGLV 1531
```

Db 1501 DKGEIOEYGA PSDLLQORGLFYTSMAKDAGLV 1531

RESULT 4
US-08-463-179A-4
Sequence 4, Application US/08463179A
Patent No. 6001563
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Dealey, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,179A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Deconti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: PQI-002CP8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-179A-4

Query Match 79.4%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTWNTSNDFPKCFONTLVLVWPCFYLMACFPFFLYLSRH 60
DB 1 MALRGFCSADGSDPLMDMNTWNTSNDFPKCFONTLVLVWPCFYLMACFPFFLYLSRH 60
QY 61 DRGVYQMTPLNKTALGFLIMTVCMADLFYSFWERSRGIFLAPVFLVSTLTIGITTLA 120
DB 61 DRGVYQMTPLNKTALGFLIMTVCMADLFYSFWERSRGIFLAPVFLVSTLTIGITTLA 120
QY 121 TPLIQERRRGVSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFFVYFS 180
DB 121 TPLIQERRRGVSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLFRDITFFVYFS 180
QY 181 LLILIQVLSCFSDRSPLFSETIHDNPCESSASFLSRITFWWTGILVNGYRPLEGSD 240
DB 181 LLILIQVLSCFSDRSPLFSETIHDNPCESSASFLSRITFWWTGILVNGYRPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNWKKECAKTRQPVKVYSSKDPQPKSSKVDANEVEAL 300

Db 241 LMSLNKEDTSEQVVPVLVKNWKKECAKTRQPVKVYSSKDPQPKSSKVDANEVEAL 300
QY 301 IVKSPQKKNPISLFXLYKTFGPFYFLMSFFPKA1HDLMSGQILKLIKFNVDTPAD 360
DB 301 IVKSPQKKNPISLFXLYKTFGPFYFLMSFFPKA1HDLMSGQILKLIKFNVDTPAD 360
QY 361 MCGFPTVLLFVNAQIOTVLHOFPHCFVSGNRKTAIVGAYRKALVTNSARKSTV 420
DB 361 MCGFPTVLLFVNAQIOTVLHOFPHCFVSGNRKTAIVGAYRKALVTNSARKSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVIALYLMNLGPSVLAVAWMLVAVN 480
DB 421 GEIVNLSVDAQRFMDLATYINMIWSAPLOVIALYLMNLGPSVLAVAWMLVAVN 480
QY 481 AVAMKTKTYQVAMHMSKONRIKLMEIINGILVILKYAMELAPKDTALROBELKVLK 540
DB 481 AVAMKTKTYQVAMHMSKONRIKLMEIINGILVILKYAMELAPKDTALROBELKVLK 540
QY 541 KSAVLSAVGTFVWCTPFLVALCTFAVYVITDENNILDQTAFLVALFNLREPILIP 600
DB 541 KSAVLSAVGTFVWCTPFLVALCTFAVYVITDENNILDQTAFLVALFNLREPILIP 600
QY 601 MVISSIVQASVSLKRLRIFLSHELEBDSIERRPVKGCGTNSITVRNATFTWARSDDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHELEBDSIERRPVKGCGTNSITVRNATFTWARSDDPT 660
QY 661 LNCITTSIPBGLVAVVGVGCGKSSLSLALAEKMRKGVHAIKGSVAVVPOAWIOND 720
DB 661 LNCITTSIPBGLVAVVGVGCGKSSLSLALAEKMRKGVHAIKGSVAVVPOAWIOND 720
QY 721 SLRENILFGQLEEPYRSTYQACALLPDEILIPSGDRITGKGVNLSCGQQRVSILAR 780
DB 721 SLRENILFGQLEEPYRSTYQACALLPDEILIPSGDRITGKGVNLSCGQQRVSILAR 780
QY 781 AYSNADIIYFDPLSAVDHVGKHI FENVIGRGKMLKNKTRILVTHSMYSYLPQVDYIIV 840
DB 781 AYSNADIIYFDPLSAVDHVGKHI FENVIGRGKMLKNKTRILVTHSMYSYLPQVDYIIV 840
QY 841 MSGGKISEMGSYQELLARDGAPAEFLRTYASTBOEDAEENGTVGSGPKKAKOMENG 900
DB 841 MSGGKISEMGSYQELLARDGAPAEFLRTYASTBOEDAEENGTVGSGPKKAKOMENG 900
QY 901 LVTDSAGKQORQLSSSSSYSGDISRHNSHTALQRAEAKKETWLTMEADKAQTGVKL 960
DB 901 LVTDSAGKQORQLSSSSSYSGDISRHNSHTALQRAEAKKETWLTMEADKAQTGVKL 960
QY 961 SVYWDYMKAIQLFISFLIFLPMCNHVSALASNYWLSLMTDDPIVNGTOEHTKVRLSVYG 1020
DB 961 SVYWDYMKAIQLFISFLIFLPMCNHVSALASNYWLSLMTDDPIVNGTOEHTKVRLSVYG 1020
QY 1021 ALGISOGIAVFGYSMAVSIQGLIASRCLHVDLHSLIRSPMSFFERTPSGNLVNRSKEL 1080
DB 1021 ALGISOGIAVFGYSMAVSIQGLIASRCLHVDLHSLIRSPMSFFERTPSGNLVNRSKEL 1080
QY 1081 DTVDSMIPEVIKMFMSLFNVIACIVILATPIAA1IIPPLGLIYFFVQRFVASSROL 1140
DB 1081 DTVDSMIPEVIKMFMSLFNVIACIVILATPIAA1IIPPLGLIYFFVQRFVASSROL 1140
QY 1141 KRLSVSRSPVYSHFNETLIGSVIRAFPEOEERFHOSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLSVSRSPVYSHFNETLIGSVIRAFPEOEERFHOSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLCEVGNCTVLPALFAVISRHSLSAGLVGSYSLOVTTYLVNMLVRMSSEMETNIYA 1260
DB 1201 VRLCEVGNCTVLPALFAVISRHSLSAGLVGSYSLOVTTYLVNMLVRMSSEMETNIYA 1260
QY 1261 VERLKEYSERKAPMOIOETAPSSWPQVGRVERFNKYCLRYREDLDVYLRIHNTVING 1320
DB 1261 VERLKEYSERKAPMOIOETAPSSWPQVGRVERFNKYCLRYREDLDVYLRIHNTVING 1320
QY 1321 EKVGIVRGAGKSSLTGLFRINESAGE1I1IDGINAKIGLHDLRFKTIIPQDPVLV 1380

Db 1321 EKVGIVGTGAGKSSLTIGLFRINESAGCEIIDGINIAKIGLHDLRKKITITIPDDPVLF 1380
QY 1381 SGLSRMLNDPFSQYSDSEEWTSLELAHLKDVSALPDXLDECAEGENLSVGOROLVCL 1440
Db 1381 SGLSRMLNDPFSQYSDSEEWTSLELAHLKDVSALPDXLDECAEGENLSVGOROLVCL 1440
QY 1441 ARALLRKTIIVLEATAVAVLETDLLQSTIRTOFEDCTVLTIAHRLNTIMDYTRIVL 1500
Db 1441 ARALLRKTIIVLEATAVAVLETDLLQSTIRTOFEDCTVLTIAHRLNTIMDYTRIVL 1500
QY 1501 DKGEIOEGAPSDLLQORGLFYMAKDAGLV 1531
Db 1501 DKGEIOEGAPSDLLQORGLFYMAKDAGLV 1531

RESULT 5
US-08-461-384B-4
Sequence 4, Application US/08461384B
Patent No. 6025473
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,384B
FILING DATE: 05-JUN-95
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01547
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-384B-4

Query Match 79.4%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADSDPLMDNMVNTMTNSNPDFTKCFONTVLVWVPCFYLMACFFEFYLYLSRH 60
Db 1 MALRGFCSADSDPLMDNMVNTMTNSNPDFTKCFONTVLVWVPCFYLMACFFEFYLYLSRH 60
QY 61 DRGIQMTPLANKTKTALGFLIMVCMADLFYSFMRSGRTFLAPVFLVSPLLGITTLLA 120
Db 61 DRGIQMTPLANKTKTALGFLIMVCMADLFYSFMRSGRTFLAPVFLVSPLLGITTLLA 120

Db 61 DRGIQMTPLANKTKTALGFLIMVCMADLFYSFMRSGRTFLAPVFLVSPLLGITTLLA 120
QY 121 TFLIOLERRKGVSSGIMLTFMWLVAVCALAIRSKIMTALKEDAQVDLFRDITFYVYFS 180
Db 121 TFLIOLERRKGVSSGIMLTFMWLVAVCALAIRSKIMTALKEDAQVDLFRDITFYVYFS 180
QY 181 LLILQVLSCFSDRSLPFSSTIHDNCPSSSASFSLRITFWMTTGLIVNGYROPLEGSD 240
Db 181 LLILQVLSCFSDRSLPFSSTIHDNCPSSSASFSLRITFWMTTGLIVNGYROPLEGSD 240
QY 241 LMSLNKEDTSEOVVAVLVKMKKECAKTRKOPVYVSSSDPAPKSSSVDAEVEAL 300
Db 241 LMSLNKEDTSEOVVAVLVKMKKECAKTRKOPVYVSSSDPAPKSSSVDAEVEAL 300
QY 301 IVKSPCKENPSSLFKVLKYKTFGPFYFLMSFFPKAIHDLMSFGPDLKLLIKFVNDTRAPD 360
Db 301 IVKSPCKENPSSLFKVLKYKTFGPFYFLMSFFPKAIHDLMSFGPDLKLLIKFVNDTRAPD 360
QY 361 WQGYFTVLLFVTRACLOTVLVHOYFHI CFVSGMKRIKTAVIGAVYRKALVITNSARKSTV 420
Db 361 WQGYFTVLLFVTRACLOTVLVHOYFHI CFVSGMKRIKTAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLSVDAORPMDLATYINMTWSAPLOVILALYLMNLGSPVLAGVAVWVLMVAVN 480
Db 421 GEIVNLSVDAORPMDLATYINMTWSAPLOVILALYLMNLGSPVLAGVAVWVLMVAVN 480
QY 481 AVMAKTKTYQVAHMSKDNRIKLMEILNGIKVLKLYAMELAFKDKVYLAIRQELKVLK 540
Db 481 AVMAKTKTYQVAHMSKDNRIKLMEILNGIKVLKLYAMELAFKDKVYLAIRQELKVLK 540
QY 541 KSAVLSAVGTFVWCYFPELVALCTFAVYVITDENNIIDAOYAVSLAFNILREPLNLP 600
Db 541 KSAVLSAVGTFVWCYFPELVALCTFAVYVITDENNIIDAOYAVSLAFNILREPLNLP 600
QY 601 MVLSIVQASVSLKRLIFLSHELEPDSIERRPVKDGGTNSITVNAATFTWARSPPPT 660
Db 601 MVLSIVQASVSLKRLIFLSHELEPDSIERRPVKDGGTNSITVNAATFTWARSPPPT 660
QY 661 LMGITFSIPGALVAVVGVCGGKSSLSLALLAEMDVEGHVAIKGSVAVVPQAMIQND 720
Db 661 LMGITFSIPGALVAVVGVCGGKSSLSLALLAEMDVEGHVAIKGSVAVVPQAMIQND 720
QY 721 SLRENIIFGCOLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSCGQKORVSLAR 780
Db 721 SLRENIIFGCOLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSCGQKORVSLAR 780
QY 781 AVYSNADIYLPDPLSAVDHAGKHIFENVIGPKMKNTKTRILVTHSMGYLPQVDVIV 840
Db 781 AVYSNADIYLPDPLSAVDHAGKHIFENVIGPKMKNTKTRILVTHSMGYLPQVDVIV 840
QY 841 MSGGKISEMSYQELARDAFAEFLRTYASTEQOEDAENGVTGVSGPKEAKOMENGM 900
Db 841 MSGGKISEMSYQELARDAFAEFLRTYASTEQOEDAENGVTGVSGPKEAKOMENGM 900
QY 901 LVYDSAGKOLQROLSSSSSYSGDISRRHNSYABLQYAKAEKERTKLMLEADKAQTGVYKL 960
Db 901 LVYDSAGKOLQROLSSSSSYSGDISRRHNSYABLQYAKAEKERTKLMLEADKAQTGVYKL 960
QY 961 SVYWDYKKAIGLPSLSTIFLMCNHVSALASNYWLSLWTDPPVNGTOHTKRLSVYG 1020
Db 961 SVYWDYKKAIGLPSLSTIFLMCNHVSALASNYWLSLWTDPPVNGTOHTKRLSVYG 1020
QY 1021 ALGISOGIAVFGYMAVSGIGILASRCGLHYDLHSILSRPMSFFERPPSGNLVNRFSKEL 1080
Db 1021 ALGISOGIAVFGYMAVSGIGILASRCGLHYDLHSILSRPMSFFERPPSGNLVNRFSKEL 1080
QY 1081 DTVDSMIPVYIKMFMSLFENVIGACIVILLATPAAIIPPLGIYFVORFYVASSROL 1140
Db 1081 DTVDSMIPVYIKMFMSLFENVIGACIVILLATPAAIIPPLGIYFVORFYVASSROL 1140
QY 1141 KRLSVSRSPYVSHFNETLLGVSYTARFEOERFIHOSDKVDENQAYVPSIVANWLA 1200
Db 1141 KRLSVSRSPYVSHFNETLLGVSYTARFEOERFIHOSDKVDENQAYVPSIVANWLA 1200

QY 1201 VRLCEVGNCLVLPALFAVIRSRHSLSAGLVGSVSYSHQVTTYLNMLVRRMSEMETNIVA 1260
Db 1201 VRLCEVGNCLVLPALFAVIRSRHSLSAGLVGSVSYSHQVTTYLNMLVRRMSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVEFRNYCLARYEDLDVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVEFRNYCLARYEDLDVLRHINVTINGG 1320
QY 1321 EKVIGVGTGAGKSLTGLFRINESAGEEIIIDGINAKIGLHDLRKITIIIIQDPVLF 1380
Db 1321 EKVIGVGTGAGKSLTGLFRINESAGEEIIIDGINAKIGLHDLRKITIIIIQDPVLF 1380
QY 1381 SGLRPMNDPFSQYSDSEEWTSLSLAHLKDFVSALPDKLDHCEAGGENLSVGRQVCL 1440
Db 1381 SGLRPMNDPFSQYSDSEEWTSLSLAHLKDFVSALPDKLDHCEAGGENLSVGRQVCL 1440
QY 1441 ARALLRKTKILVLEDEATAVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500
Db 1441 ARALLRKTKILVLEDEATAVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500
QY 1501 DKGEIOEYGAUSDLLQGRGLFYMAKDAGLV 1531
Db 1501 DKGEIOEYGAUSDLLQGRGLFYMAKDAGLV 1531

RESULT 6
US-09-647-1408-19
Sequence 19, Application US/096471408
Patent No. 6803184
GENERAL INFORMATION:
APPLICANT: Fox Chase Cancer Center
APPLICANT: Krush, Gary D.
APPLICANT: Lee, Kun
APPLICANT: Belinsky, Martin G.
TITLE OF INVENTION: Bait, Lisa J.
TITLE OF INVENTION: MRP-Related ABC Transporter Encoding
FILE REFERENCE: FCCC 98-02
CURRENT APPLICATION NUMBER: US/09/647, 1408
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: PCT/US99/06644
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 60/079, 759
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/095, 153
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 19
LENGTH: 1531
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-647-1408-19

Query Match 79.4%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 181 LLLIOLVLSGFSRDSPLFSETIHDNPNCPRESSAGFLSRITFPWITGLIVRGYRQPLEGSD 240
QY 241 LMSLNKEDTSSQVVPVLYKMKKECATRQKQPVVYVSSQDPAPKSSSKVDANEVEAL 300
Db 241 LMSLNKEDTSSQVVPVLYKMKKECATRQKQPVVYVSSQDPAPKSSSKVDANEVEAL 300
QY 301 IVKSPQKEMNPSLFKVLKFTGPFYFLMSFFPKAIIHDLMMFSGPOILKLLIKFVNDTPAD 360
Db 301 IVKSPQKEMNPSLFKVLKFTGPFYFLMSFFPKAIIHDLMMFSGPOILKLLIKFVNDTPAD 360
QY 361 WQGYFTVLLFVTAQLOTVLHQYFHI CFVSGMBIKTAVIGAVYRKALVITNSARKSSTV 420
Db 361 WQGYFTVLLFVTAQLOTVLHQYFHI CFVSGMBIKTAVIGAVYRKALVITNSARKSSTV 420
QY 421 GEIVNLSVDAQRMDLATYINMTWSAPLOYITALLYLMLNGSVYLAGVAVMTLAMPVN 480
Db 421 GEIVNLSVDAQRMDLATYINMTWSAPLOYITALLYLMLNGSVYLAGVAVMTLAMPVN 480
QY 481 AVAMAKTKTYQVAMKSKDNRIKLMELINGIKYKLYAMELAKDKVLAIROBELKYLK 540
Db 481 AVAMAKTKTYQVAMKSKDNRIKLMELINGIKYKLYAMELAKDKVLAIROBELKYLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITIDENNIIDAQTA FVSLALFNILREPLNLP 600
Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITIDENNIIDAQTA FVSLALFNILREPLNLP 600
QY 601 MVISSIVQASVSLKRLIFLSHELEPDSIERRPVKQGGGNSITVENATFTWASDPT 660
Db 601 MVISSIVQASVSLKRLIFLSHELEPDSIERRPVKQGGGNSITVENATFTWASDPT 660
QY 661 LMGITFSPGALVAVVGVGCGKSSLSALLAEMDVGBVALKGSVAVYPOQAMQND 720
Db 661 LMGITFSPGALVAVVGVGCGKSSLSALLAEMDVGBVALKGSVAVYPOQAMQND 720
QY 721 SLRENILFGQLEBPYRSVIOACALLPDEILPSGRTIEGKXVNLSSGOKORVSLAR 780
Db 721 SLRENILFGQLEBPYRSVIOACALLPDEILPSGRTIEGKXVNLSSGOKORVSLAR 780
QY 781 AVYSNADIYLPDDPLSAVDHVGKHI FENVYIGPKMLKNKTRILVTHSMSTLPOVDYIIV 840
Db 781 AVYSNADIYLPDDPLSAVDHVGKHI FENVYIGPKMLKNKTRILVTHSMSTLPOVDYIIV 840
QY 841 MSGGKISEMSYQELLARDGAPAEFLTYASTBQEOAENGVYSGPGEAQOMENG 900
Db 841 MSGGKISEMSYQELLARDGAPAEFLTYASTBQEOAENGVYSGPGEAQOMENG 900
QY 901 LVYDSAGKQORQSSSSSYSGDISRHNETABLOKAEAKKEFTWKLEADKQOTGVKL 960
Db 901 LVYDSAGKQORQSSSSSYSGDISRHNETABLOKAEAKKEFTWKLEADKQOTGVKL 960
QY 961 SVYWDYWKALGLFISFLSIFLPMCNHVSALASNWYSLWTDPIVNGTOHTKYRLSVYG 1020
Db 961 SVYWDYWKALGLFISFLSIFLPMCNHVSALASNWYSLWTDPIVNGTOHTKYRLSVYG 1020
QY 1021 ALGISOSIAVPGYSMAVNSIGIILASRCLAHYDLHSIIRSMSPFERPSSGNLVNRFSEL 1080
Db 1021 ALGISOSIAVPGYSMAVNSIGIILASRCLAHYDLHSIIRSMSPFERPSSGNLVNRFSEL 1080
QY 1081 DTYDSMIPFVIKMFMSLFPVIGACIYIILATPILAIIPPLGIYFFVORFYAASSROL 1140
Db 1081 DTYDSMIPFVIKMFMSLFPVIGACIYIILATPILAIIPPLGIYFFVORFYAASSROL 1140
QY 1141 KRLESVRSRSPVYSHFNETLLGVSIVIRAFEOERFIHQSDLKVDENQKAYVPSIVANRWLA 1200
Db 1141 KRLESVRSRSPVYSHFNETLLGVSIVIRAFEOERFIHQSDLKVDENQKAYVPSIVANRWLA 1200
QY 1201 VRLCEVGNCLVLPALFAVIRSRHSLSAGLVGSVSYSHQVTTYLNMLVRRMSEMETNIVA 1260
Db 1201 VRLCEVGNCLVLPALFAVIRSRHSLSAGLVGSVSYSHQVTTYLNMLVRRMSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVEFRNYCLARYEDLDVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVEFRNYCLARYEDLDVLRHINVTINGG 1320

QY 1321 EKVGIVGRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLDLREKTIITIPQDPVL 1380
DB 1321 EKVGIVGRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLDLREKTIITIPQDPVL 1380
QY 1381 SGSLRMLDPPSOYSDERWMTSLBLAHKDPVSALPDLHCEHCAEGENLSVGQOLVCL 1440
DB 1381 SGSLRMLDPPSOYSDERWMTSLBLAHKDPVSALPDLHCEHCAEGENLSVGQOLVCL 1440
QY 1441 ARALLRKTIILVLEBATAVLETDLLOSTIRTOFEDCTVLTIAHRLNTIMDYTRVVL 1500
DB 1441 ARALLRKTIILVLEBATAVLETDLLOSTIRTOFEDCTVLTIAHRLNTIMDYTRVVL 1500
QY 1501 DKGEIOEYGAPEDDLQORGLFYMAKADAGLV 1531
DB 1501 DKGEIOEYGAPEDDLQORGLFYMAKADAGLV 1531

RESULT 7
US-08-141-893-2
Sequence 2, Application US/08141893
Patent No. 5489519
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Deele, Roger G.
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, SUITE 510
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/141,893
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923; 08/029,340
FILING DATE: 27-OCT-1992; 8-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: PQI-002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5149
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-141-893-2

Query Match 79.3%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCADSGDPLMDNNTWNTNSPDPTKCFONTLVWVPCFYLMACFPFYLYSRH 60
DB 1 MALRGFCADSGDPLMDNNTWNTNSPDPTKCFONTLVWVPCFYLMACFPFYLYSRH 60
QY 61 DRGYOMPLNKTALGFLIMIVCMADLFYSFMRSGIPLAPVFLVSPPTLIGITTLA 120
DB 61 DRGYOMPLNKTALGFLIMIVCMADLFYSFMRSGIPLAPVFLVSPPTLIGITTLA 120

QY 121 TFLIQLERRRGVSSGIMLTFWLVAVCALAIIIRSKMTLAKEDAVDLFRDITFYVFS 180
DB 121 TFLIQLERRRGVSSGIMLTFWLVAVCALAIIIRSKMTLAKEDAVDLFRDITFYVFS 180
QY 181 LLLIQVLSCSPSRPLFSEITHDPNCPRESSASFISRIFFWMTGILVGYRQPLGSD 240
DB 181 LLLIQVLSCSPSRPLFSEITHDPNCPRESSASFISRIFFWMTGILVGYRQPLGSD 240
QY 241 LMSLNKEDTSEQVVPVVKMKKECAKTRKQPVVYVSSKDPAQPKSSKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVVKMKKECAKTRKQPVVYVSSKDPAQPKSSKVDANEVEAL 300
QY 301 IVSPQKEMNPSLFKVLKTFGPFYFLMSFFPKAIIHDMFMSGPILKLIKFNVDTPA 360
DB 301 IVSPQKEMNPSLFKVLKTFGPFYFLMSFFPKAIIHDMFMSGPILKLIKFNVDTPA 360
QY 361 WQGFYTVLLFVTAQCTLVLHOYFHIQVSGMRKTAIVGAVYRKALVITNSARKSTV 420
DB 361 WQGFYTVLLFVTAQCTLVLHOYFHIQVSGMRKTAIVGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVIALYLLMLNIGPSVLGAVAVMLVAVN 480
DB 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVIALYLLMLNIGPSVLGAVAVMLVAVN 480
QY 481 AVMAKTKTYQVAHMKSKNRIKLMNEILNGIKVLKYANELAFKDYLAIRQELKVLK 540
DB 481 AVMAKTKTYQVAHMKSKNRIKLMNEILNGIKVLKYANELAFKDYLAIRQELKVLK 540
QY 541 KSAVLSAVGFTWCPTPELVALCTPAVYVITDENNIIIDAQTAFLSLFNLIRPLNIR 600
DB 541 KSAVLSAVGFTWCPTPELVALCTPAVYVITDENNIIIDAQTAFLSLFNLIRPLNIR 600
QY 601 MVISIYQASVSLKRLRIFLSHELEBPSIERRPVKDGGTNSITVNAATFTWARSDEPT 660
DB 601 MVISIYQASVSLKRLRIFLSHELEBPSIERRPVKDGGTNSITVNAATFTWARSDEPT 660
QY 661 LMGITPESIPEGALVAVVGVQCGKSLISALLAEMDYEGHVAIKGSVAVYVQQAQOND 720
DB 661 LMGITPESIPEGALVAVVGVQCGKSLISALLAEMDYEGHVAIKGSVAVYVQQAQOND 720
QY 721 SLRENIILFGCOLLEPPYRSVYQACALLPDELILPSGRTEIGEYVNLSSGQKRVSLAR 780
DB 721 SLRENIILFGCOLLEPPYRSVYQACALLPDELILPSGRTEIGEYVNLSSGQKRVSLAR 780
QY 781 AVYENADIIYLEDPLSAVDAHVGKIFENYIGPKMKKNKRIILVTHSMSTLPQVDYIIV 840
DB 781 AVYENADIIYLEDPLSAVDAHVGKIFENYIGPKMKKNKRIILVTHSMSTLPQVDYIIV 840
QY 841 MSGKISMSGSYQELRLARDGAFAEFLRTYASTEQEDAENGVTGVSQPKAKQEMNGM 900
DB 841 MSGKISMSGSYQELRLARDGAFAEFLRTYASTEQEDAENGVTGVSQPKAKQEMNGM 900
QY 901 LVTHSAGKOLOROLSSSSSYSGDISRHNSIAELQKAEKKEFTWKLMEADKATQCYKL 960
DB 901 LVTHSAGKOLOROLSSSSSYSGDISRHNSIAELQKAEKKEFTWKLMEADKATQCYKL 960
QY 961 SVYWDYKAIKGLFISFISIFLFMCNHSALASNYWLSMTDPIVNGQOETHKRLSVYG 1020
DB 961 SVYWDYKAIKGLFISFISIFLFMCNHSALASNYWLSMTDPIVNGQOETHKRLSVYG 1020
QY 1021 ALGISOGIAVFGYMAVSGIGIILASRCLHVDLLHSILSPWSPFERTPSGNLVNRFSKEL 1080
DB 1021 ALGISOGIAVFGYMAVSGIGIILASRCLHVDLLHSILSPWSPFERTPSGNLVNRFSKEL 1080
QY 1081 DTVDSMIPEVVKMGMSLFNVIGACIYIILATPIAIIIPPLGLIYFFVQGFYVASSQOL 1140
DB 1081 DTVDSMIPEVVKMGMSLFNVIGACIYIILATPIAIIIPPLGLIYFFVQGFYVASSQOL 1140
QY 1141 KRLSEVSRSPVYSHFNETLIGSVYIRAFEBEORFIHOSDLKVDENOKRAYYSIVANRWLA 1200
DB 1141 KRLSEVSRSPVYSHFNETLIGSVYIRAFEBEORFIHOSDLKVDENOKRAYYSIVANRWLA 1200
QY 1201 VRLCEVGNICIVLFAALPAVISRHSLSAGLVGLSVSYSLQVTTYLNMILVRMSSEMETNIVA 1260

```

Db 1201 VRLECVGNCIVLFAALFVIVSHSHSAGLVGSVSYSLQVTTYLMVLVMSSEMETNIVA 1260
Qy 1261 VERLKEVSETEKAWQIOETAPPSNPQGVVERVNTCLARYRELDVLRHINTINGG 1320
Db 1261 VERLKEVSETEKAWQIOETAPPSNPQGVVERVNTCLARYRELDVLRHINTINGG 1320
Qy 1321 EKVGVGTGAGKSLTGLFPRINSAGEIITIDGINAKIGLHDLRKITITIPDDPLF 1380
Db 1321 EKVGVGTGAGKSLTGLFPRINSAGEIITIDGINAKIGLHDLRKITITIPDDPLF 1380
Qy 1381 SGLSRMNDPFSQYSDSEVWTSLELAHLKDFVSALPDKLDEHCAEGENLSVGQRYVCL 1440
Db 1381 SGLSRMNDPFSQYSDSEVWTSLELAHLKDFVSALPDKLDEHCAEGENLSVGQRYVCL 1440
Qy 1441 ARALLRKTIIVLDEATAVADLETDDLIOSTIRIQFEDCVTLTAHRLNTIMDYTRIVL 1500
Db 1441 ARALLRKTIIVLDEATAVADLETDDLIOSTIRIQFEDCVTLTAHRLNTIMDYTRIVL 1500
Qy 1501 DKGEIOEGAPSDLLQQRGLFYSMAKXAGLV 1531
Db 1501 DKGEIOEGAPSDLLQQRGLFYSMAKXAGLV 1531

RESULT 8
US-08-463-092B-2
; Sequence 2, Application US/08463092B
; Patent No. 5766880
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deesley, Roger G.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,092B
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: Q1546
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6653

```

```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-092B-2

Query Match 79.3%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MALRFGSADGSDPLMDMNTVMTNSDPFTCPONTVLVWPFCYLMACFPFFLYLSRH 60
Db 1 MALRFGSADGSDPLMDMNTVMTNSDPFTCPONTVLVWPFCYLMACFPFFLYLSRH 60
Qy 61 DRGYIOMTPLNKTALGFLMIYCMADLFYSFERSRGIFLAVFLVSPTLGITTLA 120
Db 61 DRGYIOMTPLNKTALGFLMIYCMADLFYSFERSRGIFLAVFLVSPTLGITTLA 120
Qy 121 TFLIQLERRKGVSSGIMLTFWLVALCALAIRSKIMTALKEDAOVDLPFDITFYYS 180
Db 121 TFLIQLERRKGVSSGIMLTFWLVALCALAIRSKIMTALKEDAOVDLPFDITFYYS 180
Qy 181 LLLIQVLSCPSDPSLPSFTIHDPNCPRESSASFISRTFTWMTGLIVRGYRPLEGSD 240
Db 181 LLLIQVLSCPSDPSLPSFTIHDPNCPRESSASFISRTFTWMTGLIVRGYRPLEGSD 240
Qy 241 LMSLNKEDTSEQVVPVLVKKWKKECATRKQPVKVYSSKDPAPKSSKVDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLVKKWKKECATRKQPVKVYSSKDPAPKSSKVDANEVEAL 300
Qy 301 IVKSPQKEMNSLFEKVLKTFGPVFLMSPEFKAIHDLMSGPOILKILIKFVNDTRAPD 360
Db 301 IVKSPQKEMNSLFEKVLKTFGPVFLMSPEFKAIHDLMSGPOILKILIKFVNDTRAPD 360
Qy 361 WQGYFYTVLLFVTAQCTLVLAHQYFHCFSQGMKIKTAVIGAVYRKALVITNSARKSTV 420
Db 361 WQGYFYTVLLFVTAQCTLVLAHQYFHCFSQGMKIKTAVIGAVYRKALVITNSARKSTV 420
Qy 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVITLALYLLMLNGPSVLAGVAVMLVAVN 480
Db 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVITLALYLLMLNGPSVLAGVAVMLVAVN 480
Qy 481 AVMAKTKTYQVAHMSKDNRIKLMNEILINGIKVLKYLAMELAKDRLVALIROELKVLK 540
Db 481 AVMAKTKTYQVAHMSKDNRIKLMNEILINGIKVLKYLAMELAKDRLVALIROELKVLK 540
Qy 541 KSAVLSAVGTFTWCTPFVALCTFAVYVITDENNIIIDAQTAFFVSLAFNLRPPLILP 600
Db 541 KSAVLSAVGTFTWCTPFVALCTFAVYVITDENNIIIDAQTAFFVSLAFNLRPPLILP 600
Qy 601 MVISIVQASVSLKRLIFLSHELEPDSIERPVKGGGNTSTTVNATFTMARSPPT 660
Db 601 MVISIVQASVSLKRLIFLSHELEPDSIERPVKGGGNTSTTVNATFTMARSPPT 660
Qy 661 LMGITFSPGALVAVVGVGCGKSLLSALLAEMDRVEGHVAKGSVAVVPOQAWIQND 720
Db 661 LMGITFSPGALVAVVGVGCGKSLLSALLAEMDRVEGHVAKGSVAVVPOQAWIQND 720
Qy 721 SLRENILFGQLEBPYRSYIOACALLPDLIELPSGRTIEGKGVNLSSGQKQSVLSAR 780
Db 721 SLRENILFGQLEBPYRSYIOACALLPDLIELPSGRTIEGKGVNLSSGQKQSVLSAR 780
Qy 781 AVYSNADIYFDDPELSAVDAVGHKIPENYIGPKMLKNTRILVITSMYSVLPQVDYIIV 840
Db 781 AVYSNADIYFDDPELSAVDAVGHKIPENYIGPKMLKNTRILVITSMYSVLPQVDYIIV 840
Qy 841 MSGGKISEMGSYQELARDAFAEFLTYASTROEDQAEENGVTGSGPGKEAKOMENGM 900
Db 841 MSGGKISEMGSYQELARDAFAEFLTYASTROEDQAEENGVTGSGPGKEAKOMENGM 900
Qy 901 LVYDSAKQIQORQLSSSSSYSGDISRHNSHTAELQKAEAKKEFWKLMKADKQGTGVKL 960

```

Db 901 LVTSAGAOQOROSSSSSSSSGSDISRHNSTAELOKAKEETKMEADKAQOGVKL 960
Qy 961 SVYDYMKAIGLFLSFLSIFLMCHNVSALASNYLWLTMTDDPIYNGIOETHKVLSTYVG 1020
Db 961 SVYDYMKAIGLFLSFLSIFLMCHNVSALASNYLWLTMTDDPIYNGIOETHKVLSTYVG 1020
Qy 1021 ALGISQGIADVGYMAVSIIGLILSRCLHVDLHLSILSPMSFFERTSPGNLVNRSKEL 1080
Db 1021 ALGISQGIADVGYMAVSIIGLILSRCLHVDLHLSILSPMSFFERTSPGNLVNRSKEL 1080
Qy 1081 DTVDSMTPEVTKMFGSLFNVIAGACIVILLATPIAIIIPPLGLIYFVQRPYVASSROL 1140
Db 1081 DTVDSMTPEVTKMFGSLFNVIAGACIVILLATPIAIIIPPLGLIYFVQRPYVASSROL 1140
Qy 1141 KRLSVSRSPYVSHNETILGVSVIRAFEBQERFIHOSDLKVDENQKAYPSIVANRMLA 1200
Db 1141 KRLSVSRSPYVSHNETILGVSVIRAFEBQERFIHOSDLKVDENQKAYPSIVANRMLA 1200
Qy 1201 VRLBCVNCIYVPAALFAVISRHSLSAGLVGLSVSLSQVTTYLNMVLVMSSEMETNTIVA 1260
Db 1201 VRLBCVNCIYVPAALFAVISRHSLSAGLVGLSVSLSQVTTYLNMVLVMSSEMETNTIVA 1260
Qy 1261 VERLKEYETEKEMWQIOETRPSPSPQVGRVEFRNYCLARYEDLPVLRHINVTINGG 1320
Db 1261 VERLKEYETEKEMWQIOETRPSPSPQVGRVEFRNYCLARYEDLPVLRHINVTINGG 1320
Qy 1321 EKVGVIGRTGAGKSLTGLFRINSAGEIITIDGINIAKIGLHDLRFKITIIPDDPVLF 1380
Db 1321 EKVGVIGRTGAGKSLTGLFRINSAGEIITIDGINIAKIGLHDLRFKITIIPDDPVLF 1380
Qy 1381 SGSLRMLNDPSPQSYDEEVMTSLELAHLKDFVSALPDKLDHECAGENGELSVGQROLVCL 1440
Db 1381 SGSLRMLNDPSPQSYDEEVMTSLELAHLKDFVSALPDKLDHECAGENGELSVGQROLVCL 1440
Qy 1441 ARALLKRTKIYLDATAVVDLETDDLIQSTIRTOFEEDCTVLTIAHRLNTIMDYRIVVL 1500
Db 1441 ARALLKRTKIYLDATAVVDLETDDLIQSTIRTOFEEDCTVLTIAHRLNTIMDYRIVVL 1500
Qy 1501 DKGEIOEYGAPESDLQORGLFYSMADAGLV 1531
Db 1501 DKGEIOEYGAPESDLQORGLFYSMADAGLV 1531

RESULT 9
US-08-462-109A-2
; Sequence 2, Application US/08462109A
; Patent No. 5882875
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deleley, Roger G.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,109A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992

APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: P01-002CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-2

Query Match 79.3%; Score 7849; DB 1; Length 1531;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MALRGFCGADSDPLMDNVTWNTSNPDFTYCFONTVLVWPCEYLMACPEFFLYLSRH 60
Db 1 MALRGFCGADSDPLMDNVTWNTSNPDFTYCFONTVLVWPCEYLMACPEFFLYLSRH 60
Qy 61 DRGYIOMPLNKTATGALGFLIMICWADLFYSFWERSGIFLAVFLVSPILLGTTTLA 120
Db 61 DRGYIOMPLNKTATGALGFLIMICWADLFYSFWERSGIFLAVFLVSPILLGTTTLA 120
Qy 121 TFLIOLERRKVOSSGIMLTFMLVALCALALIRSKIMTALKEDAOVDLFFDITFYVFS 180
Db 121 TFLIOLERRKVOSSGIMLTFMLVALCALALIRSKIMTALKEDAOVDLFFDITFYVFS 180
Qy 181 LLLIQLVLISCFSDRSPLFSEFIHDNPPCESSASFLSHITTWITGLIVRGYROPLEGSD 240
Db 181 LLLIQLVLISCFSDRSPLFSEFIHDNPPCESSASFLSHITTWITGLIVRGYROPLEGSD 240
Qy 241 LMSLNKERTSQVVPVLYKMKKCECAKTRKQPVVYVSSKDPAPKESSTVDANEVEAL 300
Db 241 LMSLNKERTSQVVPVLYKMKKCECAKTRKQPVVYVSSKDPAPKESSTVDANEVEAL 300
Qy 301 IVKSPQKEMNPSLFKVLKTFGPFYLSFFPKAIIHDLMPGSPOLIKLIRFVNDTKAPD 360
Db 301 IVKSPQKEMNPSLFKVLKTFGPFYLSFFPKAIIHDLMPGSPOLIKLIRFVNDTKAPD 360
Qy 361 WQGYFYTLVFLVTAQTLVHGYFHHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420
Db 361 WQGYFYTLVFLVTAQTLVHGYFHHICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420
Qy 421 GEIVNLMVDQRPMDLATYINMTSAPLOVITLALYLMLNLGSPVLGAVAVMLMVVN 480
Db 421 GEIVNLMVDQRPMDLATYINMTSAPLOVITLALYLMLNLGSPVLGAVAVMLMVVN 480
Qy 481 AVMAKTKTYOVAAHKSNDRIKLNNEILNGIKVUKLYAMELAFDQVLAIROBELKYLK 540
Db 481 AVMAKTKTYOVAAHKSNDRIKLNNEILNGIKVUKLYAMELAFDQVLAIROBELKYLK 540
Qy 541 KSAVLSAVGTFTWCTPPLVALCTPAVYVITDENNIIDAQTAFFVALFNILRPLNLTLP 600
Db 541 KSAVLSAVGTFTWCTPPLVALCTPAVYVITDENNIIDAQTAFFVALFNILRPLNLTLP 600
Qy 601 MYTSIVQSVSLKRLRIFLSHEELPDSTERRPPYKDGCTNSTVVRNATFTWASDPT 660
Db 601 MYTSIVQSVSLKRLRIFLSHEELPDSTERRPPYKDGCTNSTVVRNATFTWASDPT 660
Qy 661 LNGITFESIPEGALVAVGVQVGGKSLLSALLAEVDYEGHAYIKGSVAAYVQQAQWIND 720
Db 661 LNGITFESIPEGALVAVGVQVGGKSLLSALLAEVDYEGHAYIKGSVAAYVQQAQWIND 720

QY 721 SLRENIIFGCOLLEBYRYSVIAQACALLPDLLEIPSGDRTEIGKGVNLGGOKORVSLAR 780
DB 721 SLRENIIFGCOLLEBYRYSVIAQACALLPDLLEIPSGDRTEIGKGVNLGGOKORVSLAR 780
QY 781 AVYSNADIIYLFDDPLISAVDAHVXKHIIFENVIGPKMKLNKTRILLVTHSMSTYLPQVDVIV 840
DB 781 AVYSNADIIYLFDDPLISAVDAHVXKHIIFENVIGPKMKLNKTRILLVTHSMSTYLPQVDVIV 840
QY 841 MSGKISMGSVOELLARDGAFAELRTYASTEDQDAEENGVCVSGPGEAKQEMNGM 900
DB 841 MSGKISMGSVOELLARDGAFAELRTYASTEDQDAEENGVCVSGPGEAKQEMNGM 900
QY 901 LVTSAGKOLOROLSSSSSSSGDISRHNSTAELOKAKKEETKLMADKAOTGOYKL 960
DB 901 LVTSAGKOLOROLSSSSSSSGDISRHNSTAELOKAKKEETKLMADKAOTGOYKL 960
QY 961 SVYDYMKAIGLFIISFLSIFLPMCNHVSALASNTWLSMTDDPIVNGTOEHTKVLSTYG 1020
DB 961 SVYDYMKAIGLFIISFLSIFLPMCNHVSALASNTWLSMTDDPIVNGTOEHTKVLSTYG 1020
QY 1021 ALGISOGIAVFGYGNVAVSIGIILASRCLHVDLHSLILSPMSFPERTPSSGULVNRFSKEL 1080
DB 1021 ALGISOGIAVFGYGNVAVSIGIILASRCLHVDLHSLILSPMSFPERTPSSGULVNRFSKEL 1080
QY 1081 DTVDSMIPEVIKMFMSGLFENVIGACIVILLATPIAIIIPPLGLIYFVQRFVYASSROL 1140
DB 1081 DTVDSMIPEVIKMFMSGLFENVIGACIVILLATPIAIIIPPLGLIYFVQRFVYASSROL 1140
QY 1141 KRLESVSRSPYVSHNETLLGVSVIRAFEOERFIHOSDLKVDENOKAYVESIVANRWLA 1200
DB 1141 KRLESVSRSPYVSHNETLLGVSVIRAFEOERFIHOSDLKVDENOKAYVESIVANRWLA 1200
QY 1201 VRLCEVNGCIYLFALPAVIRHSISAGLVLSYSYISQVTTYANLVRKSEMETNIVA 1260
DB 1201 VRLCEVNGCIYLFALPAVIRHSISAGLVLSYSYISQVTTYANLVRKSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMOIOETAPPSMPQVGRVERBNCLYRREDLDVLRHINVTNGG 1320
DB 1261 VERLKEYSETEKEAPMOIOETAPPSMPQVGRVERBNCLYRREDLDVLRHINVTNGG 1320
QY 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIDIINIAKIGLHDLREKTIIPQDPVLF 1380
DB 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIDIINIAKIGLHDLREKTIIPQDPVLF 1380
QY 1381 SGRIRNMLDPPSOYSDEBWTSLBLAKDVSALPKDLHCEBAGGENLSVGRQOLVCL 1440
DB 1381 SGRIRNMLDPPSOYSDEBWTSLBLAKDVSALPKDLHCEBAGGENLSVGRQOLVCL 1440
QY 1441 ABALLRKTIIIVLEBATAVDLETTDDLIQSTIRTOFERCTVLTJAHRLNTIMDYTRIVL 1500
DB 1441 ABALLRKTIIIVLEBATAVDLETTDDLIQSTIRTOFERCTVLTJAHRLNTIMDYTRIVL 1500
QY 1501 DKGEIOEGAPSDLLQORGLFYSNAKDAGLV 1531
DB 1501 DKGEIOEGAPSDLLQORGLFYSNAKDAGLV 1531

RESULT 10
US-08-460-907B-2
; Sequence 2, Application us/08460907B
; Patent No. 5891724
; GENERAL INFORMATION:
; APPLICANT: Deeley, Roger G.
; APPLICANT: Cole, Susan P. C.
; TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PARTEO RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA

ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: us/08/460, 907B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966, 923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029, 340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141, 893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407, 207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01551
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-907B-2

Query Match 79.3%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTWNTSNDFPKCFONTLVVWPCFYLMACEPFYLYSRH 60
DB 1 MALRGFCSADGSDPLMDMNTWNTSNDFPKCFONTLVVWPCFYLMACEPFYLYSRH 60
QY 61 DRGIVQMTPLNKTKTALGFLIMVYCADLFYSFWERSRGIFLAPVFLVSPFLIGITTLA 120
DB 61 DRGIVQMTPLNKTKTALGFLIMVYCADLFYSFWERSRGIFLAPVFLVSPFLIGITTLA 120
QY 121 TFLIQLERRRGVSSGIMLTFMLVALCALILRSKIMTLAKEDAOVDLFRDITFYVYFS 180
DB 121 TFLIQLERRRGVSSGIMLTFMLVALCALILRSKIMTLAKEDAOVDLFRDITFYVYFS 180
QY 181 LLILIQVLSCFSDRSPLFSEITHDPNCPRESSASFLSRIITFWMITGLIVAGYRQPLEGSD 240
DB 181 LLILIQVLSCFSDRSPLFSEITHDPNCPRESSASFLSRIITFWMITGLIVAGYRQPLEGSD 240
QY 241 LMSLNKEDTSEOVVPLVVKWKKECAKTRQPVKVTYSSDPAQPKSSRYDANEVEAL 300
DB 241 LMSLNKEDTSEOVVPLVVKWKKECAKTRQPVKVTYSSDPAQPKSSRYDANEVEAL 300
QY 301 IVKSPQKEMNPSLFKVLYKTFGPFILMSFFPKAIHDLMEFSGQILELLIKFVNDTAPD 360
DB 301 IVKSPQKEMNPSLFKVLYKTFGPFILMSFFPKAIHDLMEFSGQILELLIKFVNDTAPD 360
QY 361 WGGFYTVLLFVTRACIOTLVLHQYFHIQVSGMRKIKTAIVGAYRKALVITNSARKSSTV 420
DB 361 WGGFYTVLLFVTRACIOTLVLHQYFHIQVSGMRKIKTAIVGAYRKALVITNSARKSSTV 420

```

QY 421 GEIYNLMSVDAQRPMDLATYINMTWSAPLOVITLALYLMNLGSPVLGAVAVMTLMPVN 480
DB 421 GEIYNLMSVDAQRFMDLATYINMTWSAPLOVITLALYLMNLGSPVLGAVAVMTLMPVN 480
QY 481 AVMMAMKTKTYOVAMHKSNDRIKLMNEILNGIKVLKLYAMELAFDKVLAIROEELKYLK 540
DB 481 AVMMAMKTKTYOVAMHKSNDRIKLMNEILNGIKVLKLYAMELAFDKVLAIROEELKYLK 540
QY 541 KSAVLSAVGFTWVCTPFLVALCTPAVYVITDENNIIDAQTAFVSLAFNLRPEPLNLP 600
DB 541 KSAVLSAVGFTWVCTPFLVALCTPAVYVITDENNIIDAQTAFVSLAFNLRPEPLNLP 600
QY 601 MVISIVQASVSLKRLRIFLSHELEPDSIERPPYKGGGNSITVRNATTTWASDPPT 660
DB 601 MVISIVQASVSLKRLRIFLSHELEPDSIERPPYKGGGNSITVRNATTTWASDPPT 660
QY 661 LINGTFPSIPGALVAVVGVQCGKSLLSALLAEMDKYGHVATIGSAVAYVPOQAMIOND 720
DB 661 LINGTFPSIPGALVAVVGVQCGKSLLSALLAEMDKYGHVATIGSAVAYVPOQAMIOND 720
QY 721 SLRENIJFCQLEBEPYRSVIOACALLPDLBILPBGDRTEIGEKGVNLGGQKORVSLAR 780
DB 721 SLRENIJFCQLEBEPYRSVIOACALLPDLBILPBGDRTEIGEKGVNLGGQKORVSLAR 780
QY 781 AVYSNADIIYLPDDPLSANDAHYKHI FENVIGPKMKLKNKTRILVTHSKSYLPQVDVITV 840
DB 781 AVYSNADIIYLPDDPLSANDAHYKHI FENVIGPKMKLKNKTRILVTHSKSYLPQVDVITV 840
QY 841 MSGGKISMGYSYOELLARDGAPAEFLPTAYASTEBOADAEENGVTGVSQPKAKOMENGM 900
DB 841 MSGGKISMGYSYOELLARDGAPAEFLPTAYASTEBOADAEENGVTGVSQPKAKOMENGM 900
QY 901 LVTDGAKOLOROLSSSSSSSGDISRHHNSTALOKEAKEETKLMADKAOCTQOYKL 960
DB 901 LVTDGAKOLOROLSSSSSSSGDISRHHNSTALOKEAKEETKLMADKAOCTQOYKL 960
QY 961 SVYWDYMKAIGLFISFLSIFLPMCHVSAALASNYWLSMTDDPIVNGTQERTKVLSTYVG 1020
DB 961 SVYWDYMKAIGLFISFLSIFLPMCHVSAALASNYWLSMTDDPIVNGTQERTKVLSTYVG 1020
QY 1021 ALGTSOGIAVFGYSNAVSIIGLILASRCLAVDLHSILRSPMSFEERTSGNLYNFSKEL 1080
DB 1021 ALGTSOGIAVFGYSNAVSIIGLILASRCLAVDLHSILRSPMSFEERTSGNLYNFSKEL 1080
QY 1081 DTVDSMIEPVIKMFMSGLFENVIGACIVILLATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
DB 1081 DTVDSMIEPVIKMFMSGLFENVIGACIVILLATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
QY 1141 KRLESVSRSPYVSHNETLLGVSVIRAFEBQERFIHQSDLKVDENQKAYPSIVANRWLA 1200
DB 1141 KRLESVSRSPYVSHNETLLGVSVIRAFEBQERFIHQSDLKVDENQKAYPSIVANRWLA 1200
QY 1201 VRLGVGNCIVLPALFPAVISHSLSAGLVGVSYSIQVTTYLWMLVMSSEMETNTIVA 1260
DB 1201 VRLGVGNCIVLPALFPAVISHSLSAGLVGVSYSIQVTTYLWMLVMSSEMETNTIVA 1260
QY 1261 VERLKEVETEKAPMOIQETAPPSMPQVGEFRNCLARREDLDFVLRLINVTINGG 1320
DB 1261 VERLKEVETEKAPMOIQETAPPSMPQVGEFRNCLARREDLDFVLRLINVTINGG 1320
QY 1321 EKVGVIGRTGAGKSSLTGLFRINESABGEIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
DB 1321 EKVGVIGRTGAGKSSLTGLFRINESABGEIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
QY 1381 SSGSLRMINDPBSQYDEDEWMTSLBLAHLKDFVSALPKLDHFCAGGENTLSTGQOYVCL 1440
DB 1381 SSGSLRMINDPBSQYDEDEWMTSLBLAHLKDFVSALPKLDHFCAGGENTLSTGQOYVCL 1440
QY 1441 ARALIRKTKILVLDATAVADDETDLIOSTIRTOPEDCVTLIAHRLNTIMDYTRVIVL 1500
DB 1441 ARALIRKTKILVLDATAVADDETDLIOSTIRTOPEDCVTLIAHRLNTIMDYTRVIVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMADAGLV 1531

```

```

DB 1501 DKGEIOEYGAPSDLLQORGLFYSMADAGLV 1531

RESULT 11
US-08-463-179A-2
; Sequence 2, Application US/08463179A
; Patent No. 6001563
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,179A
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConli, Giulio A. Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: POI-002CP8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-463-179A-2

Query Match 79.3%; Score 7849; DB 2; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCGADSDPPLMDNNVTWNTSNPDTKCFQNTVLAWVPCFYLMAKCPFFYFLYLSRH 60
DB 1 MALRGFCGADSDPPLMDNNVTWNTSNPDTKCFQNTVLAWVPCFYLMAKCPFFYFLYLSRH 60
QY 61 DRGYIOMTPLNKTXTALGFLIMIVCMADLFYSFMRSGIFLAPVFLVSPTLIGITTLA 120
DB 61 DRGYIOMTPLNKTXTALGFLIMIVCMADLFYSFMRSGIFLAPVFLVSPTLIGITTLA 120
QY 121 TFLIOLERRKGVQSSGIMLTWLVALVCAALILNSKINTALKEDAQVDLFRDITFYVYS 180
DB 121 TFLIOLERRKGVQSSGIMLTWLVALVCAALILNSKINTALKEDAQVDLFRDITFYVYS 180
QY 181 LLTIOLVSCFSDSDSPLESETIHDNPNCPRESSASFLSITFMWITGLIVRGROPLBGSD 240
DB 181 LLTIOLVSCFSDSDSPLESETIHDNPNCPRESSASFLSITFMWITGLIVRGROPLBGSD 240

```

QY 241 LMSLNKEDTSBQVVPVLVKNMKKCAKTRKQPVKVYSSKDPAPQKSSKYDANEVEAL 300
DB 241 LMSLNKEDTSBQVVPVLVKNMKKCAKTRKQPVKVYSSKDPAPQKSSKYDANEVEAL 300
QY 301 IVKSPQKEMNPSLRKVLKTKGPPYFLMSFFPKAHDLMESGPOILKLLIFVNDTKAPD 360
DB 301 IVKSPQKEMNPSLRKVLKTKGPPYFLMSFFPKAHDLMESGPOILKLLIFVNDTKAPD 360
QY 361 MGSFYFVTLFVTLCTLOTLVHOYFHI CFSVGMRKTAIVAGAVRKALVITNSARKSTV 420
DB 361 MGSFYFVTLFVTLCTLOTLVHOYFHI CFSVGMRKTAIVAGAVRKALVITNSARKSTV 420
QY 421 GEIYNLMSVDAQRMDLATYINMIWSAPLQVYLALYLLMLNGPSVLAGVAVMLAMPVN 480
DB 421 GEIYNLMSVDAQRMDLATYINMIWSAPLQVYLALYLLMLNGPSVLAGVAVMLAMPVN 480
QY 481 AVMAKTKTYOVAMHMSKDNRIKLMNEILNGIKYVKLYAMELAFKDYLAIRQELKYLK 540
DB 481 AVMAKTKTYOVAMHMSKDNRIKLMNEILNGIKYVKLYAMELAFKDYLAIRQELKYLK 540
QY 541 KSAVLAVGTFTWCTPFLVLCCTFAVYVTTIDENNILDAQTAFLSLFNLRRPMLP 600
DB 541 KSAVLAVGTFTWCTPFLVLCCTFAVYVTTIDENNILDAQTAFLSLFNLRRPMLP 600
QY 601 MVLSSIVQASVSLKRLIFLSHEELEPDSIERRPVKOGGNSITVRNATFTMARSPPT 660
DB 601 MVLSSIVQASVSLKRLIFLSHEELEPDSIERRPVKOGGNSITVRNATFTMARSPPT 660
QY 661 LMGITFESIPBGALVAVVGVGCGKSLSLALLAEMDKYGHVAIKGSVAVVPOQAMIOND 720
DB 661 LMGITFESIPBGALVAVVGVGCGKSLSLALLAEMDKYGHVAIKGSVAVVPOQAMIOND 720
QY 721 SLRENILPGCOLBEPYRSVIOACALLPDLIELPSGRTEGEGVNLSSGQKRVSLAR 780
DB 721 SLRENILPGCOLBEPYRSVIOACALLPDLIELPSGRTEGEGVNLSSGQKRVSLAR 780
QY 781 AVYSNADIYLFDDPLSAVDAVGHIFENVIGPKMLKNKRLIVTMSMSYLPDQVYIIV 840
DB 781 AVYSNADIYLFDDPLSAVDAVGHIFENVIGPKMLKNKRLIVTMSMSYLPDQVYIIV 840
QY 841 MSGGKISEMSGYOELLARDGAFBLRTYASTEODAEENGVTGSGPKAKOMENGM 900
DB 841 MSGGKISEMSGYOELLARDGAFBLRTYASTEODAEENGVTGSGPKAKOMENGM 900
QY 901 LYTDSACKOJQROULSSSSSSSGDISRHNSYAELOKABAKKEEYWKLEADKATQGVKL 960
DB 901 LYTDSACKOJQROULSSSSSSSGDISRHNSYAELOKABAKKEEYWKLEADKATQGVKL 960
QY 961 SVYWDYKAIQGLFISFISIFLFCNHYSAALASNYWLSLMTDPIVNGTOEHTKRLSVYG 1020
DB 961 SVYWDYKAIQGLFISFISIFLFCNHYSAALASNYWLSLMTDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISOGIAVFGSMAVSIIGIILASRCLHVDLHSLIRSPMSFEERTPSGLVNRFSKEL 1080
DB 1021 ALGISOGIAVFGSMAVSIIGIILASRCLHVDLHSLIRSPMSFEERTPSGLVNRFSKEL 1080
QY 1081 DTYDSMTPEVITKMGSLFNVTGACIYIILATPIAIIIPPLGIYTFVQRFYASSROL 1140
DB 1081 DTYDSMTPEVITKMGSLFNVTGACIYIILATPIAIIIPPLGIYTFVQRFYASSROL 1140
QY 1141 KRLSVSRSPVSHFNETLGVSVIRAFEBEORFIHOSDLKVDENOKAYPSIYANMWA 1200
DB 1141 KRLSVSRSPVSHFNETLGVSVIRAFEBEORFIHOSDLKVDENOKAYPSIYANMWA 1200
QY 1201 VRLCEVNCIYLFALPAVISRHSLSAGLVGLSVYSLOVTTYLNMVLRSSSEMETIVA 1260
DB 1201 VRLCEVNCIYLFALPAVISRHSLSAGLVGLSVYSLOVTTYLNMVLRSSSEMETIVA 1260
QY 1261 VERKESESETEKAPMOIOETAPSSSWPQVGRVAFNRYCLARYRDLDPVLRHINVTNGG 1320
DB 1261 VERKESESETEKAPMOIOETAPSSSWPQVGRVAFNRYCLARYRDLDPVLRHINVTNGG 1320
QY 1321 EKVGIVGRTAGKSSLTLGLFRINESABGEIILIDGINIAKIGLHDLRFKLTIIIPQDPVLF 1380

DB 1321 EKVGIVGRTAGKSSLTLGLFRINESABGEIILIDGINIAKIGLHDLRFKLTIIIPQDPVLF 1380
QY 1381 SSGSLRMLDPPSQYSDEEWTSLLELAHLKDFVSLPKDLHECAGGEMNSVSGOROLVCL 1440
DB 1381 SSGSLRMLDPPSQYSDEEWTSLLELAHLKDFVSLPKDLHECAGGEMNSVSGOROLVCL 1440
QY 1441 ARALLRKTILVDEATAVADLETDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVYL 1500
DB 1441 ARALLRKTILVDEATAVADLETDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVYL 1500
QY 1501 DKGEIOEYGAPSDLIOQRGLFYSMAXDAGLV 1531
DB 1501 DKGEIOEYGAPSDLIOQRGLFYSMAXDAGLV 1531
RESULT 12
US-08-461-384B-2
Sequence 2, Application US/08461384B
Patent No. 6025473
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: PARTO RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,384B
FILING DATE: 05-JUN-95
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Mlernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01547
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-384B-2
Query Match 79.3%; Score 7849; DB 2; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MALRGFCSADGSDPLMDMNTNTSNDFTKCFQNTLVVWPCFYIMACPPFYLYISRH 60
DB 1 MALRGFCSADGSDPLMDMNTNTSNDFTKCFQNTLVVWPCFYIMACPPFYLYISRH 60
QY 61 DRGYIQMTPLNKTALGLFLWIVCWADLFYSFWERSRGIFLAPVFLVSPYTLTGITTLA 120

```

Db      61 DRGIQMTPLNKRTALGFLIMVCMADLFYSFMRSGRIFLAPVFLSPILLGTTLLA 120
Qy      121 TFLQLERRKGVSSGIMLTFMLVALVCAALILRSKIMTALKEPDQVDFPDITFYVES 180
Db      121 TFLQLERRKGVSSGIMLTFMLVALVCAALILRSKIMTALKEPDQVDFPDITFYVES 180
Qy      181 LLLIQLVLSGSDSPLESETIHDNPCESSASFLSRITFMWITGLIVRGYROPLEGSD 240
Db      181 LLLIQLVLSGSDSPLESETIHDNPCESSASFLSRITFMWITGLIVRGYROPLEGSD 240
Qy      241 LMSLNKEDTSQVVPVLYVNMKKKCAKTRKQPVKVVYSSKDPAPQKSSKVDANEVVAL 300
Db      241 LMSLNKEDTSQVVPVLYVNMKKKCAKTRKQPVKVVYSSKDPAPQKSSKVDANEVVAL 300
Qy      301 IVKSPQKEMNPSLRFVLKTFGFPYFLMSFFPKAIDLMFMFSGPOLIKLLIFVNDTKAPD 360
Db      301 IVKSPQKEMNPSLRFVLKTFGFPYFLMSFFPKAIDLMFMFSGPOLIKLLIFVNDTKAPD 360
Qy      361 WQGFYTVLTFVTAQTLVLHOYFPHICFVSGMRIKTAVIGAVERKALVITNSARKSSTV 420
Db      361 WQGFYTVLTFVTAQTLVLHOYFPHICFVSGMRIKTAVIGAVERKALVITNSARKSSTV 420
Qy      421 GEIYNLMSVDQRFMDLATYINMISAPLOVILAYLMLNGPSVLGAVAVMVLMEVFN 480
Db      421 GEIYNLMSVDQRFMDLATYINMISAPLOVILAYLMLNGPSVLGAVAVMVLMEVFN 480
Qy      481 AVMAKTKTYOVANHKSQDNRIKLMNEILNGIKVLKLYAMELAFQDKYLAIROBELKYLK 540
Db      481 AVMAKTKTYOVANHKSQDNRIKLMNEILNGIKVLKLYAMELAFQDKYLAIROBELKYLK 540
Qy      541 KSAVLASAVGTFPTWCTPFLVALCTFAVYVITDENNILDQAFAVSLAFNLLRPPLNLP 600
Db      541 KSAVLASAVGTFPTWCTPFLVALCTFAVYVITDENNILDQAFAVSLAFNLLRPPLNLP 600
Qy      601 MVISSIVASVSLKRLRIFLSHEBELPDSIERRPYKQGGGNSITVRNATFTWASDPPT 660
Db      601 MVISSIVASVSLKRLRIFLSHEBELPDSIERRPYKQGGGNSITVRNATFTWASDPPT 660
Qy      661 LMGITFSLPEGALVAVVQVQGGKSLLSALLAEMDKVEGHVAKGSAVYVPOQMIQND 720
Db      661 LMGITFSLPEGALVAVVQVQGGKSLLSALLAEMDKVEGHVAKGSAVYVPOQMIQND 720
Qy      721 SLRENIIFGCOLLEBYRVSVOACALPDLLEILPSGDRTEIGEKNVLSGGQKQFVSJAR 780
Db      721 SLRENIIFGCOLLEBYRVSVOACALPDLLEILPSGDRTEIGEKNVLSGGQKQFVSJAR 780
Qy      781 AVYSNADIIYLFDDPLSADAHVGHKIFBNVIGPKMLKXKTRILVTHSMSTYLPQVDVILV 840
Db      781 AVYSNADIIYLFDDPLSADAHVGHKIFBNVIGPKMLKXKTRILVTHSMSTYLPQVDVILV 840
Qy      841 MSGGKISMGYSOELLARDGAFAEFLRTYASTBEOQDAEBNGVGVSGPKGKAKOMENGM 900
Db      841 MSGGKISMGYSOELLARDGAFAEFLRTYASTBEOQDAEBNGVGVSGPKGKAKOMENGM 900
Qy      901 LVTTDSAGKOLOROLSSSSSSSGDISSRHNSTAELOKAEAKTEETKMLEADKAOQGYKL 960
Db      901 LVTTDSAGKOLOROLSSSSSSSGDISSRHNSTAELOKAEAKTEETKMLEADKAOQGYKL 960
Qy      961 SVYWDYMKAIGLFISFLSIFLPMCHNVSAALASNYWLSMTDDPIYNGTQEHTRVLSYVG 1020
Db      961 SVYWDYMKAIGLFISFLSIFLPMCHNVSAALASNYWLSMTDDPIYNGTQEHTRVLSYVG 1020
Qy      1021 ALGTSOGIAVFGYSAAVSIIGLILASRCHVDLHLSILSPMSFFERTSGNLYNFSKEL 1080
Db      1021 ALGTSOGIAVFGYSAAVSIIGLILASRCHVDLHLSILSPMSFFERTSGNLYNFSKEL 1080
Qy      1081 DTVDSMIDEVIMKFMGSLFNVVIGACIVILLATPILAIITPILGILYFFQRFYVSSRQL 1140
Db      1081 DTVDSMIDEVIMKFMGSLFNVVIGACIVILLATPILAIITPILGILYFFQRFYVSSRQL 1140
Qy      1141 KRLSVSRSPYVSHENETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYVPSIVANRWLA 1200

```

```

Db      1141 KRLSVSRSPYVSHENETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYVPSIVANRWLA 1200
Qy      1201 VRLECVGNCTIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMVLRMSSEMETNIVA 1260
Db      1201 VRLECVGNCTIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMVLRMSSEMETNIVA 1260
Qy      1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVFNRYCIRYRDDLPVLRHINVTINGG 1320
Db      1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVFNRYCIRYRDDLPVLRHINVTINGG 1320
Qy      1321 EKVGIVRTAGKSLTLGLFRINESAEGRIIIDGINIAKIGLHDLFPKTIIPQDPVLF 1380
Db      1321 EKVGIVRTAGKSLTLGLFRINESAEGRIIIDGINIAKIGLHDLFPKTIIPQDPVLF 1380
Qy      1381 SGLSRMNLDPFSQYSDSEVWTSLELAHLKDVSALPDKLDHECAEGGENLSVGOROLVCL 1440
Db      1381 SGLSRMNLDPFSQYSDSEVWTSLELAHLKDVSALPDKLDHECAEGGENLSVGOROLVCL 1440
Qy      1441 ARALLRRTKILIVDEATPAVDLETFDDLIQSTIRFOFEDCTVLTIAHRLNTIMDYTRIVL 1500
Db      1441 ARALLRRTKILIVDEATPAVDLETFDDLIQSTIRFOFEDCTVLTIAHRLNTIMDYTRIVL 1500
Qy      1501 DKGEIOEYGAPSDLLQQRGLFYSMKADAGLV 1531
Db      1501 DKGEIOEYGAPSDLLQQRGLFYSMKADAGLV 1531

```

```

RESULT 13
US-08-407-207A-2
; Sequence 2, Application US/08407207A
; Patent No. 6063621
; GENERAL INFORMATION:
; APPLICANT: Deeley, Roger G.
; APPLICANT: Cole, Susan P. C.
; TITLE OF INVENTION: ANTIBODIES TO A MULTIDRUG RESISTANCE PROTEIN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESS: PARTO RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/407,207A
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: 01512
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-407-207A-2

```

Query Match 79.3%; Score 7849; DB 2; Length 1531;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCGADSDPLMDMNTVMTSNDPTKCFONTVLWVPCFYLMACFPPEFYLSRH 60
 DB 1 MALRGFCGADSDPLMDMNTVMTSNDPTKCFONTVLWVPCFYLMACFPPEFYLSRH 60
 QY 61 DRGCIQMTPLNKTALGFLIMTYCMADLFYSFWERSRGIFLAVFVLSPTLLGTTLLA 120
 DB 61 DRGCIQMTPLNKTALGFLIMTYCMADLFYSFWERSRGIFLAVFVLSPTLLGTTLLA 120
 QY 121 TFLQLRRKGVSSGIMLTFLMALVLCALILRSKIMTALKEDAQVDLPFDITFYVFS 180
 DB 121 TFLQLRRKGVSSGIMLTFLMALVLCALILRSKIMTALKEDAQVDLPFDITFYVFS 180
 QY 181 LLLIQLVLSGFSRSPLESETIHDNPPCESSASFLSRITPFWITGLVRSYRQPLGSD 240
 DB 181 LLLIQLVLSGFSRSPLESETIHDNPPCESSASFLSRITPFWITGLVRSYRQPLGSD 240
 QY 241 LMSLNKEDTSGQVTVLVKWKKCAKTRKOPVRYVYSSKDPAPCKSSKYDANEVEAL 300
 DB 241 LMSLNKEDTSGQVTVLVKWKKCAKTRKOPVRYVYSSKDPAPCKSSKYDANEVEAL 300
 QY 301 IVKSPKEMNPSSLFKVLKTRGPEFLMSFFPKAIDHLMFSGQILKLIKEVNDTRAPD 360
 DB 301 IVKSPKEMNPSSLFKVLKTRGPEFLMSFFPKAIDHLMFSGQILKLIKEVNDTRAPD 360
 QY 361 WOGFYFVTLFVTACLOTVLVHQYFHLCEVSGMRKTAIVAGAVRKALVITNSARKSTV 420
 DB 361 WOGFYFVTLFVTACLOTVLVHQYFHLCEVSGMRKTAIVAGAVRKALVITNSARKSTV 420
 QY 421 GEIYNLMSVDAQRFMDIATYINMIWSAPLOYIALYILMLNGSVLAGAVMLWLPVN 480
 DB 421 GEIYNLMSVDAQRFMDIATYINMIWSAPLOYIALYILMLNGSVLAGAVMLWLPVN 480
 QY 481 AVAMKTKTYOVAMKSKDNRIKLMNELLNGIKYLKYAMELAKDYVLARQELVVK 540
 DB 481 AVAMKTKTYOVAMKSKDNRIKLMNELLNGIKYLKYAMELAKDYVLARQELVVK 540
 QY 541 KSAVLASAVGTFTWCTPEFLVALCTFAVYVITDENNNILDAQTAFFSLAFNLRPELILP 600
 DB 541 KSAVLASAVGTFTWCTPEFLVALCTFAVYVITDENNNILDAQTAFFSLAFNLRPELILP 600
 QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVBNATFTMARSDPT 660
 DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVBNATFTMARSDPT 660
 QY 661 LINGITFSPGALVAVVGVGCCGKSLLSLALMMDVREGHVALKGSVAVYPOQAWI QND 720
 DB 661 LINGITFSPGALVAVVGVGCCGKSLLSLALMMDVREGHVALKGSVAVYPOQAWI QND 720
 QY 721 SLRENILFGCOLBEPYRSVYIACALPDLLEILPSGRTEIGEGVNLSSGQKORVSLAR 780
 DB 721 SLRENILFGCOLBEPYRSVYIACALPDLLEILPSGRTEIGEGVNLSSGQKORVSLAR 780
 QY 781 AVYSNADIYLPDDELASVDAHVGHIFENYIGPKMLKNKTRILLVTHSMSYLPQVDYIIV 840
 DB 781 AVYSNADIYLPDDELASVDAHVGHIFENYIGPKMLKNKTRILLVTHSMSYLPQVDYIIV 840
 QY 841 MSGGKISEMGSYQELRLARDGAFAPFLTYASTEBEQDAEENGVTGVSPPGEAAQOMENGM 900
 DB 841 MSGGKISEMGSYQELRLARDGAFAPFLTYASTEBEQDAEENGVTGVSPPGEAAQOMENGM 900
 QY 901 LVYDSAGKOLQROLSSSSSYSGDISRRHNSTAELOKAEAKKEETWKLMEADKAGTQGVKL 960
 DB 901 LVYDSAGKOLQROLSSSSSYSGDISRRHNSTAELOKAEAKKEETWKLMEADKAGTQGVKL 960
 QY 961 SVYWDYKAIAGLTFISLSTFLPMGNHYSALASNYWLSLWTDPIVNGTOEHTKRLSVYG 1020
 DB 961 SVYWDYKAIAGLTFISLSTFLPMGNHYSALASNYWLSLWTDPIVNGTOEHTKRLSVYG 1020

QY 1021 ALGISOGIAVFGYSAVVISGIIILASRCLAHVDLHSLIRSBSMSPFEPTPSGNLVNRFSEKL 1080
 DB 1021 ALGISOGIAVFGYSAVVISGIIILASRCLAHVDLHSLIRSBSMSPFEPTPSGNLVNRFSEKL 1080
 QY 1081 DTVDSMIPEVIKMEFMSLFNVIGACIVILLATPIAIIIPPLGLIYEFVORFYVASSROL 1140
 DB 1081 DTVDSMIPEVIKMEFMSLFNVIGACIVILLATPIAIIIPPLGLIYEFVORFYVASSROL 1140
 QY 1141 KRLSVRSRPSVYSHFNFTLLGVSVYIRAFEEQERTIHOSDKVDENQAYYPSIYANRWLA 1200
 DB 1141 KRLSVRSRPSVYSHFNFTLLGVSVYIRAFEEQERTIHOSDKVDENQAYYPSIYANRWLA 1200
 QY 1201 VRLCEVNCIYLPFAALFAVISRHSLSAGLVGSYSLOVTTYANMLVRSSEMETNIVA 1260
 DB 1201 VRLCEVNCIYLPFAALFAVISRHSLSAGLVGSYSLOVTTYANMLVRSSEMETNIVA 1260
 QY 1261 VERLKEYSETEKEAPWQIOETAPSSWPQYRVEFRNYCLRYREDLPVLRIHINTINGG 1320
 DB 1261 VERLKEYSETEKEAPWQIOETAPSSWPQYRVEFRNYCLRYREDLPVLRIHINTINGG 1320
 QY 1321 EKVGIVRTGAGKSSLTGLFRINESABGEIIIDGINIAKIGLHDLFKITIIIPQDVLV 1380
 DB 1321 EKVGIVRTGAGKSSLTGLFRINESABGEIIIDGINIAKIGLHDLFKITIIIPQDVLV 1380
 QY 1381 SGLRNMULDPEFSQYSDSEVWTSLELAHLKDFVSALPKDLHECAGGENTSVGOROLVCL 1440
 DB 1381 SGLRNMULDPEFSQYSDSEVWTSLELAHLKDFVSALPKDLHECAGGENTSVGOROLVCL 1440
 QY 1441 ABALLRRTKILVLDENAAVLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMOTRIVYL 1500
 DB 1441 ABALLRRTKILVLDENAAVLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMOTRIVYL 1500
 QY 1501 DKGEIOEYGA PSDLLQORGLFYSMAXDAGLV 1531
 DB 1501 DKGEIOEYGA PSDLLQORGLFYSMAXDAGLV 1531

RESULT 14
 US-08-463-092B-6
 ; Sequence 6, Application US/08463092B
 ; Patent No. 5766880
 ; GENERAL INFORMATION:
 ; APPLICANT: Cole, Susan P.C.
 ; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
 ; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PARTOE RESEARCH & DEVELOPMENT INNOVATIONS
 ; STREET: Queen's University at Kingston
 ; CITY: Kingston
 ; STATE: Ontario
 ; COUNTRY: CANADA
 ; ZIP: K7L 3N6
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/463, 092B
 ; FILING DATE: 05-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/966, 923
 ; FILING DATE: 27-OCT-1992
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/029, 340
 ; FILING DATE: 8-MAR-1993
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/141, 893

FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-092B-6

Query Match 70.7%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRGFCGADGSDPILMDNVTWNTSNPDFTKCFONTVLWVPCFYLMACFPFFYLYLSRH 60
DB 1 MALRSCFSGADGSDPILMDNVTWNTSNPDFTKCFONTVLWVPCFYLMACFPFFYLYLSRH 60
QY 61 DRGTYQMTPLNKTALGFLMLYVCMADLFYSFWRSRSGITPLAPFLVSPILLGITTLA 120
DB 61 DRGTYQMTPLNKTALGFLMLYVCMADLFYSFWRSRSGITPLAPFLVSPILLGITTLA 120
QY 121 TFLQLERKGVSSGIMLTFMLVALVCALILRSKIMTALKEDAQVDLPFDIFPYVYS 180
DB 121 TFLQLERKGVSSGIMLTFMLVALVCALILRSKIMTALKEDAQVDLPFDIFPYVYS 180
QY 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPRESSASFLSRITFWMTIGLIVRGYRPLEGSD 240
DB 181 LVLVOLVLSCFSDSCPLFSETHDNPCESSASFLSRITFWMTIGMVMVHGROLESDD 240
QY 241 LMSLNKETSQVQVPLVYKMKKECAKTRKQPVKVVYS-SKDPAPKSSKVDANEVEA 299
DB 241 LMSLNKETSQVQVPLVYKMKKECAKTRKQPVKVVYS-SKDPAPKSSKVDANEVEA 299
QY 300 LTVSPKEMKPSLPKVLKTFGPFYFLMSFPFKAIHDMFSGPOLIKLVINDTKAP 359
DB 300 LTVSPKEMKPSLPKVLKTFGPFYFLMSFPFKAIHDMFSGPOLIKLVINDTKAP 359
QY 360 DMOGFYTVLLFVTAQIOTLVLMHQYFHI CFVSGMRKTAIVAGAVYRKALVITNSARKSST 419
DB 360 DMOGFYTVLLFVTAQIOTLVLMHQYFHI CFVSGMRKTAIVAGAVYRKALVITNSARKSST 419
QY 420 VGEIVNLMSVDAQRFMDLATTYINMIWASAPLOYITALLYLMLNGPSVLAGVAVVLMVPL 479
DB 420 VGEIVNLMSVDAQRFMDLATTYINMIWASAPLOYITALLYLMLNGPSVLAGVAVVLMVPL 479
QY 480 NAWAMKTKTYQVAMHMSKDNRIKLMNEILINGIKVLKLYAMELAKVDVLAIROBELVYL 539
DB 480 NAWAMKTKTYQVAMHMSKDNRIKLMNEILINGIKVLKLYAMELAKVDVLAIROBELVYL 539
QY 540 KKSAYLSAVGTFTWCTPFLVALCTFAVYVITIDENNIIIDAQTAFFVSLALFNILREPLNLT 599
DB 540 KKSAYLSAVGTFTWCTPFLVALCTFAVYVITIDENNIIIDAQTAFFVSLALFNILREPLNLT 599
QY 600 PMVTSISIVQASVSLKRLRIPLSHEELPDSIERRPVKOGGSGNSTVNNATFTMARSDPP 659
DB 600 PMVTSISIVQASVSLKRLRIPLSHEELPDSIERRPVKOGGSGNSTVNNATFTMARSDPP 659
QY 660 TLNGITTSIPGALVAVVGVGCGKSSLLSALLAMDMKVEGVALKGSVAVYPOQAMION 719
DB 660 TLNGITTSIPGALVAVVGVGCGKSSLLSALLAMDMKVEGVALKGSVAVYPOQAMION 719

QY 720 DSLRENTLFGCQLEBPYRSYVIAQCALLPDLEIILPSCDRTEIGEKVNLSSGQKQRYSLA 779
DB 720 DSLRENTLFGHPIQENYKAMKMEACALLPDLEIILPSCDRTEIGEKVNLSSGQKQRYSLA 779
QY 780 RAVYSNADITLFPDPLSAVDAAVHGKHI FENVIGKGMKKNKTRILYTHMSYLPQVPII 839
DB 780 RAVYSNADITLFPDPLSAVDAAVHGKHI FENVIGKGMKKNKTRILYTHMSYLPQVPII 839
QY 840 VMSGKISMGYSYOELLARDGAFAEPLRTYASTEOBDAEBNGTVGSGPKKAKOMENG 899
DB 840 VMSGKISMGYSYOELLARDGAFAEPLRTYASTEOBDAEBNGTVGSGPKKAKOMENG 899
QY 900 MLVTDSAGKOLQROLSGSSSYSGDISHHNSTAELQALAKKETWKLMEADKQOTQYK 959
DB 900 MLVTDSAGKOLQROLSGSSSYSGDISHHNSTAELQALAKKETWKLMEADKQOTQYK 959
QY 960 LSVYWDYMKALGLFISFLSLFLPFCNHNVSALASNYWLSLMTDDP-PYNGOEHKRVLSV 1018
DB 960 LSVYWDYMKALGLFISFLSLFLPFCNHNVSALASNYWLSLMTDDP-PYNGOEHKRVLSV 1018
QY 1019 YGALGISQGLAVFGYSNAVSIGGILASRCUVDLHSLNSPMSFFERTPSGNLVNRFSK 1078
DB 1019 YGALGISQGLAVFGYSNAVSIGGILASRCUVDLHSLNSPMSFFERTPSGNLVNRFSK 1078
QY 1079 ELDTVDNMTPBYVIMFNGSLFPNVTGACIVILLATPIAIIIPPLGLYFFVQRYVASSR 1138
DB 1079 ELDTVDNMTPBYVIMFNGSLFPNVTGACIVILLATPIAIIIPPLGLYFFVQRYVASSR 1138
QY 1139 QLRKLESVSRSPVYSHNETLLGVSIRAFEBQERFHOSDLKYDENOKAYPSIVANRW 1198
DB 1139 QLRKLESVSRSPVYSHNETLLGVSIRAFEBQERFHOSDLKYDENOKAYPSIVANRW 1198
QY 1199 LAVRLKVCNGCIVLFAALFAVISRHSLSAGLVGSLVSQVTTYLNLVMSSEMETNI 1258
DB 1199 LAVRLKVCNGCIVLFAALFAVISRHSLSAGLVGSLVSQVTTYLNLVMSSEMETNI 1258
QY 1255 LAVRLKVCNGCIVLFAALFAVISRHSLSAGLVGSLVSQVTTYLNLVMSSEMETNI 1255
DB 1255 LAVRLKVCNGCIVLFAALFAVISRHSLSAGLVGSLVSQVTTYLNLVMSSEMETNI 1255
QY 1256 VAVRLKXESTKEAPMOIOETAPPSWPQVGVBERNYCLARYEDLDVYLRHINVTIN 1318
DB 1256 VAVRLKXESTKEAPMOIOETAPPSWPQVGVBERNYCLARYEDLDVYLRHINVTIN 1318
QY 1319 GGEKVGIIVGTGAKSSLTGLPRINSABEIIIDGINAKIGLHDLRFKITTIIPODPV 1378
DB 1319 GGEKVGIIVGTGAKSSLTGLPRINSABEIIIDGINAKIGLHDLRFKITTIIPODPV 1378
QY 1379 LFGSGLKMNLDPFQYDDEEYWTSLFLAHLKDFFYSALPDKLDHCEAGGENTLVSQORQV 1438
DB 1379 LFGSGLKMNLDPFQYDDEEYWTSLFLAHLKDFFYSALPDKLDHCEAGGENTLVSQORQV 1438
QY 1436 CLARALIRKTIILVDEATAVADLETDNLIQSTIRTOFEDCTVLTIAHRLNTIMDYRVI 1495
DB 1436 CLARALIRKTIILVDEATAVADLETDNLIQSTIRTOFEDCTVLTIAHRLNTIMDYRVI 1495
QY 1499 VLDKGEIOEYKABESDILQORGLFYSMAKDAGLV 1531
DB 1499 VLDKGEIOEYKABESDILQORGLFYSMAKDAGLV 1531
QY 1531 VLDKGEIOEYKABESDILQORGLFYSMAKDAGLV 1531
DB 1531 VLDKGEIOEYKABESDILQORGLFYSMAKDAGLV 1531

RESULT 15
US-08-462-109A-6
Sequence 6, Application US/08462109A
Patent No. 5882875
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Dealey, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA

```

? ZIP: 02109
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: ASCII text
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/462,109A
? FILING DATE:
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/966,923
? FILING DATE: 27-OCT-1992
? APPLICATION NUMBER: 08/029,340
? FILING DATE: 8-MAR-1993
? APPLICATION NUMBER: 08/141,893
? FILING DATE: 26-OCT-1993
? APPLICATION NUMBER: 08/407,207
? FILING DATE: 20-MAR-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: DeConti, Giulio A. Jr.
? REGISTRATION NUMBER: 31,503
? REFERENCE/DOCKET NUMBER: PQ1-002CP4
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (617) 227-7400
? TELEFAX: (617) 227-5941
? INFORMATION FOR SRO ID NO: 6:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1528 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
US-08-462-109A-6

```

```

Query Match 70.7%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

```

```

QY 1 MARGFSSADGSDPLMNMNTNTNSNDFTKCFONTYLVVPCFYLLACPFYLYSRH 60
DB 1 MARGFSSADGSDPLMNMNTNTNSNDFTKCFONTYLVVPCFYLLACPFYLYSRH 60
QY 61 DRGVIQMTPLNKTALGFLMTIVCMADLFYSPMERSRGIFLAPVFLVSPFLGITTLLA 120
DB 61 DRGVIQMTPLNKTALGFLMTIVCMADLFYSPMERSRGIFLAPVFLVSPFLGITTLLA 120
QY 121 TFLIQERRRGVSSGIMLTFWLVALCALALRSKIMTALKEBQVDFRDIITFYVYS 180
DB 121 TFLIQERRRGVSSGIMLTFWLVALCALALRSKIMTALKEBQVDFRDIITFYVYS 180
QY 181 LLLIQVLSCFSQDRSPLESETIHDNRCPESSASFLSRITFWMITGLIYNGYQPLEGSD 240
DB 181 LLLIQVLSCFSQDRSPLESETIHDNRCPESSASFLSRITFWMITGLIYNGYQPLEGSD 240
QY 241 LMSLNKEDTSEOVVPLVKNMKKECAKTRKQPYKVYS-SKDPAPQESSKVDANEVEA 299
DB 241 LMSLNKEDTSEOVVPLVKNMKKECAKTRKQPYKVYS-SKDPAPQESSKVDANEVEA 299
QY 300 LIYKSPQKWNPSLFLVLYKTFPGPYFLMSFFKAIHDLNMFSGPOLIKLILKEVNDTKAP 359
DB 300 LIYKSPQKWNPSLFLVLYKTFPGPYFLMSFFKAIHDLNMFSGPOLIKLILKEVNDTKAP 359
QY 360 DMGGYFTVLLPTTACIQTLVHOYFHTICVSGMRITKAVIGAVYRKALVITNSARKST 419
DB 360 DMGGYFTVLLPTTACIQTLVHOYFHTICVSGMRITKAVIGAVYRKALVITNSARKST 419
QY 420 VGSIVNLMSVDAORFMDLATYINMISAPQVILATLWLNIGPSVLAGAVVAVMLVPV 479
DB 420 VGSIVNLMSVDAORFMDLATYINMISAPQVILATLWLNIGPSVLAGAVVAVMLVPV 479
QY 480 NAYMAKTKTYQVAHMKSKDNRIKLANEILNGIKVLKLYWAMELAFKDKVLAIROBELKVL 539
DB 480 NAYMAKTKTYQVAHMKSKDNRIKLANEILNGIKVLKLYWAMELAFKDKVLAIROBELKVL 539
QY 481 NAYMAKTKTYQVAHMKSKDNRIKLANEILNGIKVLKLYWAMELAFKDKVLAIROBELKVL 540
DB 481 NAYMAKTKTYQVAHMKSKDNRIKLANEILNGIKVLKLYWAMELAFKDKVLAIROBELKVL 540

```

```

QY 540 KKSAYLSAVGTFMTWCTPPELVALCTPAVYVITDENNTLDAQTA FVSLAFNLILRPNLIL 599
DB 541 KKSAYLSAVGTFMTWCTPPELVALCTPAVYVITDENNTLDAQTA FVSLAFNLILRPNLIL 600
QY 600 PMYISSIVQASVSLKRLRIFLSHBELEPDSIERRPVQDGGTNSITVRNATFTARSDDP 659
DB 601 PMYISSIVQASVSLKRLRIFLSHBELEPDSIERRSIKSGSG-NSITYKMTFTARSEPP 659
QY 660 TLNGITFSIPEGALVAVVGVGCGKSSLSALLAEMDKVGHVAIKGSVAVYPOQANIOW 719
DB 660 TLNGITFSIPEGALVAVVGVGCGKSSLSALLAEMDKVGHVAIKGSVAVYPOQANIOW 719
QY 720 DSLRENTLPECCQEEPPRYRIVQACALLPDLLETLPSSDRTEIEBKGNISGGQKORVSLA 779
DB 720 DSLRENTLPECCQEEPPRYRIVQACALLPDLLETLPSSDRTEIEBKGNISGGQKORVSLA 779
QY 780 RAYVSNADITLPDDPLSAVDAAHVKHIFENVIGPKMKLKKRTRILVTHSMSYLPOVDVYI 839
DB 780 RAYVSNADITLPDDPLSAVDAAHVKHIFENVIGPKMKLKKRTRILVTHSMSYLPOVDVYI 839
QY 840 VMSGKISEMGSYQELLARDGAPAEPLRTYASTEOBQDAEENGVTGSGKAKOMENG 899
DB 840 VMSGKISEMGSYQELLARDGAPAEPLRTYASTEOBQDAEENGVTGSGKAKOMENG 899
QY 900 MLVTDAGKQLOQOLSSSSSSGSDIRHNHNTAELQAKAKKEETWKLMAADKAOQOVK 959
DB 900 MLVTDAGKQLOQOLSSSSSSGSDIRHNHNTAELQAKAKKEETWKLMAADKAOQOVK 959
QY 960 LSYVMDYMKRIGLIFSLIFLPMCNHVSALASNYMLSLWTD-PIYNGOEHTKRLASV 1018
DB 956 LSYVMDYMKRIGLIFSLIFLPMCNHVSALASNYMLSLWTD-PIYNGOEHTKRLASV 1015
QY 1019 YGALGISQGIAGVGYMAVSIIGIILASRCLAVDLHLSILRSPMSFPERTPSGNLVNRFSK 1078
DB 1016 YGALGISQGIAGVGYMAVSIIGIILASRCLAVDLHLSILRSPMSFPERTPSGNLVNRFSK 1075
QY 1079 ELDTYSMTPEVIMKMGSLPNYIGACIYLATLPAAIILPPLGLIYFVQRPYVASSR 1138
DB 1076 ELDTYSMTPEVIMKMGSLPNYIGACIYLATLPAAIILPPLGLIYFVQRPYVASSR 1135
QY 1139 QLRKLSVSRSPVYSHPNETLIGSVYIRAFEEQERFIHOSDLKVDENOKAYYSIVANRW 1198
DB 1136 QLRKLSVSRSPVYSHPNETLIGSVYIRAFEEQERFIHOSDLKVDENOKAYYSIVANRW 1195
QY 1199 LAVRLCEVGNCTVLPALPAVISRHSLSAGLVGLSVSYSLOVTTYLNMLVRMSSEMETNI 1258
DB 1196 LAVRLCEVGNCTVLPALPAVISRHSLSAGLVGLSVSYSLOVTTYLNMLVRMSSEMETNI 1255
QY 1259 VAVERLKEYSETEKAPWQIOETAPPSWPQVGRVFRNYCLRIYREDLDFVLRHIVTIN 1318
DB 1256 VAVERLKEYSETEKAPWQIOETAPPSWPQVGRVFRNYCLRIYREDLDFVLRHIVTIN 1315
QY 1319 GGEKVGIVGRTGAGKSLTLGLFRINESAGEIITIDGINIAKIGHDLRPFKTIIPDDPV 1378
DB 1316 GGEKVGIVGRTGAGKSLTLGLFRINESAGEIITIDGINIAKIGHDLRPFKTIIPDDPV 1375
QY 1379 LFGSLRNMLDPESSQYDEEVTSLTSLAHLKDFVSLAPDKLHCEAGSENLVSGOROLV 1438
DB 1376 LFGSLRNMLDPESSQYDEEVTSLTSLAHLKDFVSLAPDKLHCEAGSENLVSGOROLV 1435
QY 1439 CLARALLRKTILVDEATVAVDLETDLIQSTIRPOFEDCTVLTIAHRLNTIMDYTRVY 1498
DB 1436 CLARALLRKTILVDEATVAVDLETDLIQSTIRPOFEDCTVLTIAHRLNTIMDYTRVY 1495
QY 1499 VLDKGEIQEYGAAPSDLLQORGLFYSWAKDAGLV 1531
DB 1496 VLDKGEIVREGAPSELLOQRGIFYSWAKDAGLV 1528

```

Search completed: December 15, 2005, 15:49:32
Job time : 52.0811 secs

This page is blank.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 15:05:00 ; Search time 159,832 Seconds
(without alignments)
5037.528 Million cell updates/sec

Title: US-10-665-283-1

Perfect score: 9903

Sequence: 1 MALRGFCSADGSDPLMDMNV.....RSVAVAKXKPKFSIPDLSL 1927

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9903	100.0	1927	US-10-665-283-1	Sequence 1, Appli
2	9891	99.9	1927	US-10-665-283-6	Sequence 6, Appli
3	9873	99.7	1947	US-10-665-283-8	Sequence 8, Appli
4	9734	98.3	1891	US-10-665-283-4	Sequence 4, Appli
5	7860	79.4	1531	US-09-939-853A-86	Sequence 86, Appli
6	7860	79.4	1531	US-10-667-891-6	Sequence 6, Appli
7	7860	79.4	1531	US-10-807-466-6	Sequence 6, Appli
8	7860	79.4	1531	US-10-484-577-678	Sequence 19, Appli
9	7860	79.4	1531	US-10-889-503-19	Sequence 5033, Ap
10	7860	79.4	1531	US-10-756-149-5033	Sequence 87, Appli
11	7769	78.5	1515	US-09-939-853A-87	Sequence 42, Appli
12	7763	78.1	1515	US-10-618-281-42	Sequence 1718, Ap
13	7037	71.1	1388	US-10-408-765A-1718	Sequence 88, Appli
14	7002.5	70.7	1528	US-09-939-853A-88	Sequence 63, Appli
15	6068.5	61.3	1303	US-10-618-281-63	Sequence 2, Appli
16	4866.5	49.1	1911	US-10-665-283-2	Sequence 9, Appli
17	4867.5	47.4	1875	US-10-665-283-5	Sequence 5, Appli
18	4667.5	47.1	1895	US-10-665-283-9	Sequence 1, Appli
19	4491.5	45.4	1901	US-10-665-283-10	Sequence 83, Appli
20	4491.5	45.4	1527	US-09-939-853A-83	Sequence 142, Ap
21	4491.5	45.4	1527	US-10-295-027-1342	Sequence 6, Appli
22	4487.5	45.3	1530	US-10-889-503-6	Sequence 33, Appli
23	4487.5	45.3	1530	US-10-889-503-33	Sequence 55, Appli
24	4479.5	45.2	1530	US-10-818-281-55	Sequence 85, Appli
25	4396	44.4	1522	US-09-939-853A-85	Sequence 12, Appli
26	4396	44.4	1522	US-10-665-283-12	
27	4135	41.8	1941	US-10-665-283-12	

28	4105	41.5	1961	US-10-665-283-7	Sequence 7, Appli
29	3966	40.0	1905	US-10-665-283-3	Sequence 3, Appli
30	3879.5	39.2	1548	US-10-807-466-28	Sequence 28, Appli
31	3860.5	39.0	1548	US-10-667-891-1	Sequence 1, Appli
32	3860.5	39.0	1548	US-10-807-466-1	Sequence 23, Appli
33	3846	38.8	1538	US-10-807-466-23	Sequence 25, Appli
34	3841	38.8	1538	US-10-807-466-25	Sequence 24, Appli
35	3835	38.7	1538	US-10-807-466-24	Sequence 26, Appli
36	3817.5	38.5	1539	US-10-807-466-26	Sequence 1326, A
37	3817	38.5	1591	US-11-097-143-13266	Sequence 27, Appli
38	3797	38.3	1538	US-10-807-466-27	Sequence 16, Appli
39	3779	38.2	1538	US-10-807-466-16	Sequence 18, Appli
40	3774	38.1	1538	US-10-807-466-18	Sequence 22, Appli
41	3773	38.1	1538	US-10-807-466-22	Sequence 17, Appli
42	3768	38.0	1538	US-10-807-466-17	Sequence 19, Appli
43	3754	37.9	1538	US-10-807-466-19	Sequence 20, Appli
44	3750.5	37.9	1539	US-10-807-466-20	Sequence 25348, A
45	3733	37.7	1547	US-11-097-143-25248	

ALIGNMENTS

RESULT 1									
US-10-665-283-1									
Sequence 1, Application US/10665283									
Publication No. US20050063989A1									
GENERAL INFORMATION:									
APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE									
APPLICANT: DERAND, Renaud									
APPLICANT: GARCIA, Elisabeth									
APPLICANT: PROST, Anne-Lise									
APPLICANT: REVILLAUD, Jean									
APPLICANT: VIVAUDOU, Michel									
TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREO									
FILE REFERENCE: F263US104									
CURRENT APPLICATION NUMBER: US/10/665,283									
CURRENT FILING DATE: 2003-09-22									
NUMBER OF SEQ ID NOS: 25									
SOFTWARE: PatentIn version 3.1									
SEQ ID NO 1									
LENGTH: 1927									
TYPE: PRT									
ORGANISM: Homo sapiens									
US-10-665-283-1									
Query Match									
Best Local Similarity 100.0%; Score 9903; DB 5; Length 1927;									
Matches 1927; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MALRGFCSADGSDPLMDMNVNTWNTSNDFPKCFONTLVVWPCFYIMACPFYFLYSRH	60						
DB	1	MALRGFCSADGSDPLMDMNVNTWNTSNDFPKCFONTLVVWPCFYIMACPFYFLYSRH	60						
QY	61	DRGYIOWTPNKTKTALGFLMTVCADLTFYSFWEBSRGIFLAPVFLVSTPLTGITTLA	120						
DB	61	DRGYIOWTPNKTKTALGFLMTVCADLTFYSFWEBSRGIFLAPVFLVSTPLTGITTLA	120						
QY	121	TFPIQLERRRGVSSGIMLTFWVALVCALAIIRSKIMTLAKEDAQVDLFRDITFYVFS	180						
DB	121	TFPIQLERRRGVSSGIMLTFWVALVCALAIIRSKIMTLAKEDAQVDLFRDITFYVFS	180						
QY	181	LLLIQVLSCFSDRSPFLFSETIHDPNFCPESSASFISRTIFWMTGLIVRGYQPLEGSD	240						
DB	181	LLLIQVLSCFSDRSPFLFSETIHDPNFCPESSASFISRTIFWMTGLIVRGYQPLEGSD	240						
QY	241	LWSLNKEDTSEQVVPVIVKWKKECAKTRQPKVYVSSKDPAPQPKSSSVVDANEVEAL	300						
DB	241	LWSLNKEDTSEQVVPVIVKWKKECAKTRQPKVYVSSKDPAPQPKSSSVVDANEVEAL	300						
QY	301	IYKSPQKWNPSLFKVLKTFGYPFLMSFPFKAIHDLMSGQIILKLIKFVNDTRAPD	360						
DB	301	IYKSPQKWNPSLFKVLKTFGYPFLMSFPFKAIHDLMSGQIILKLIKFVNDTRAPD	360						

QY 361 WOGFYTVLLFVTAQLOTLVLHQQYFHICFVSQMRKTAIVAGVYRKALVTINSARKSSTV 420
DB 361 WOGFYTVLLFVTAQLOTLVLHQQYFHICFVSQMRKTAIVAGVYRKALVTINSARKSSTV 420
QY 421 GEIYNLMSVDAQRMDLATYINMTWSAPLOVITLALYLMNLGSRVLAVAVMLMPVN 480
DB 421 GEIYNLMSVDAQRMDLATYINMTWSAPLOVITLALYLMNLGSRVLAVAVMLMPVN 480
QY 481 AVAMAKTKTYOVANMKSKONRIKLMNELNGIKVUKLYAMELAFKDKVLAIROBELKYLK 540
DB 481 AVAMAKTKTYOVANMKSKONRIKLMNELNGIKVUKLYAMELAFKDKVLAIROBELKYLK 540
QY 541 KSAVLSAVGFTFWCTPPLVALCTPFAVYVTTIDENNIIIDAQTAFVSALFNILRPLNLTLP 600
DB 541 KSAVLSAVGFTFWCTPPLVALCTPFAVYVTTIDENNIIIDAQTAFVSALFNILRPLNLTLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660
QY 661 LINGTFSIPREGALVAVVGVQCGKSSLSALLAEMDKVGHVAIKGSVAVYPOQAMIOND 720
DB 661 LINGTFSIPREGALVAVVGVQCGKSSLSALLAEMDKVGHVAIKGSVAVYPOQAMIOND 720
QY 721 SLRENILFEGCOLBEPYRSVIOACALLPDLLETPSGDRTEIGEKVNLSSGQOKORVSLAR 780
DB 721 SLRENILFEGCOLBEPYRSVIOACALLPDLLETPSGDRTEIGEKVNLSSGQOKORVSLAR 780
QY 781 AVYENADITYLEPDDPLSANDAVHKKHIFENVIGPKMKLNKTRILVTHSMSTLPQVDVITV 840
DB 781 AVYENADITYLEPDDPLSANDAVHKKHIFENVIGPKMKLNKTRILVTHSMSTLPQVDVITV 840
QY 841 MSGGKISEMGSYOELLARDGAFABFLRTYASTEOBDAEENGVTGVSQPKAKOMENGM 900
DB 841 MSGGKISEMGSYOELLARDGAFABFLRTYASTEOBDAEENGVTGVSQPKAKOMENGM 900
QY 901 LVTDGAKOQOROLSSSSSSSGDISRRHNSTAELOKAEKKEETWKLMEADKQOTQVYL 960
DB 901 LVTDGAKOQOROLSSSSSSSGDISRRHNSTAELOKAEKKEETWKLMEADKQOTQVYL 960
QY 961 SVYNDYMAKIGLFLISFLIFLPMCHVSAALSNYMLMTDPIYNGIOEHTKRLSYG 1020
DB 961 SVYNDYMAKIGLFLISFLIFLPMCHVSAALSNYMLMTDPIYNGIOEHTKRLSYG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLHVDLHSLKSPMSFPERTSQNLVNRFSKEL 1080
DB 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLHVDLHSLKSPMSFPERTSQNLVNRFSKEL 1080
QY 1081 DTVMSEMIPEVIKMFNGSLFENVIGACIVILLATPIAIIIPPLGLYFFVQRFYVASSBOL 1140
DB 1081 DTVMSEMIPEVIKMFNGSLFENVIGACIVILLATPIAIIIPPLGLYFFVQRFYVASSBOL 1140
QY 1141 KRLESVSSPYVSHNETLIGSVIRARPEOERFHOSDLKVDENQKAYPSIVANRWLA 1200
DB 1141 KRLESVSSPYVSHNETLIGSVIRARPEOERFHOSDLKVDENQKAYPSIVANRWLA 1200
QY 1201 VRLBVCVACIVLFAALFAVISHSLSAGLVGSLVSIVLQVTTYLWMLVMSSEMETNIVA 1260
DB 1201 VRLBVCVACIVLFAALFAVISHSLSAGLVGSLVSIVLQVTTYLWMLVMSSEMETNIVA 1260
QY 1261 VERLKEVSETEKEAPWOIETAPSPSWPOVGRVEFRNYCLARYEDLDFVLRHINVTINGG 1320
DB 1261 VERLKEVSETEKEAPWOIETAPSPSWPOVGRVEFRNYCLARYEDLDFVLRHINVTINGG 1320
QY 1321 EKVGVIGRTGAGKSSLTGLFRINSABGEIIIDGINIAKIGLHDLRFKITTIIPODPVLF 1380
DB 1321 EKVGVIGRTGAGKSSLTGLFRINSABGEIIIDGINIAKIGLHDLRFKITTIIPODPVLF 1380
QY 1381 SGSLKMNLDPPSOYDEDEWVTSLELAHLKDPVSAALPDKLDHBCAGGENTLSTGQROLVCL 1440
DB 1381 SGSLKMNLDPPSOYDEDEWVTSLELAHLKDPVSAALPDKLDHBCAGGENTLSTGQROLVCL 1440

QY 1441 ARALIRKTKILVUDEATAANDLETDLIQSTIRQPEDCVLTFAHRLNTIMDTYIVL 1500
DB 1441 ARALIRKTKILVUDEATAANDLETDLIQSTIRQPEDCVLTFAHRLNTIMDTYIVL 1500
QY 1501 DKGEIOEYGA PSDILOQRGLFYSAKADAGLVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
DB 1501 DKGEIOEYGA PSDILOQRGLFYSAKADAGLVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
QY 1561 PRYRERRRARFVSKKQNCNVAHKNIREQGRFLQDVEFTTVLDLKWPHTLIFTMSFLCSW 1620
DB 1561 PRYRERRRARFVSKKQNCNVAHKNIREQGRFLQDVEFTTVLDLKWPHTLIFTMSFLCSW 1620
QY 1621 LLEFAMWMLLAFPAHGDLAPBEGTNVPCVTSIHSSSAFLPSIEVOVITIGRGRAWTEBEP 1680
DB 1621 LLEFAMWMLLAFPAHGDLAPBEGTNVPCVTSIHSSSAFLPSIEVOVITIGRGRAWTEBEP 1680
QY 1681 LAIILIVQNVGMINAIMGCIEMKTAQAHRAEFLISKHAVITLHRGLCFMLRVG 1740
DB 1681 LAIILIVQNVGMINAIMGCIEMKTAQAHRAEFLISKHAVITLHRGLCFMLRVG 1740
QY 1741 DLKRSMTISATIHQVVRKTTSPBGEVPLHQVDIPMENGVGNGIFLVADLIYHVIDS 1800
DB 1741 DLKRSMTISATIHQVVRKTTSPBGEVPLHQVDIPMENGVGNGIFLVADLIYHVIDS 1800
QY 1801 NSPIYDLAPSDLHHQDLEIIVILEGYVETTGITTOARTSTYLADEILMGCQRFVPIVAED 1860
DB 1801 NSPIYDLAPSDLHHQDLEIIVILEGYVETTGITTOARTSTYLADEILMGCQRFVPIVAED 1860
QY 1861 GRYSVDYSKFGNTIKVPLPCTARQLEDERSILDALTLASSRGPRLKRSVAVAKAPKFS 1920
DB 1861 GRYSVDYSKFGNTIKVPLPCTARQLEDERSILDALTLASSRGPRLKRSVAVAKAPKFS 1920

QY 1921 ISPDLS 1927
DB 1921 ISPDLS 1927

RESULT 2
US-10-665-283-6
; Sequence 6, Application US/10665283
; Publication No. US20050063989A1
; GENERAL INFORMATION:
; APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
; APPLICANT: DERAND, Renaud
; APPLICANT: GARCIA, Elisabeth
; APPLICANT: PROST, Anne-Lise
; APPLICANT: REVILLON, Jean
; APPLICANT: VINAUDOU, Michel
; TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF
; FILE REFERENCE: F263US104
; CURRENT APPLICATION NUMBER: US/10/665, 283
; CURRENT FILING DATE: 2003-09-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1927
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-665-283-6

Query Match 99.9%; Score 9891; DB 5; Length 1927;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1925; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCADSDPLMDNNTWNTNSDPFTCFQNTVLYWVPCFYLMACPPPYFLYSRH 60
DB 1 MALRGFCADSDPLMDNNTWNTNSDPFTCFQNTVLYWVPCFYLMACPPPYFLYSRH 60
QY 61 DRGYOMPLUNKTKTALGFLMIVCMADLFYSFMERSGIFLAPVFLVSPTLGITTLIA 120
DB 61 DRGYOMPLUNKTKTALGFLMIVCMADLFYSFMERSGIFLAPVFLVSPTLGITTLIA 120
QY 121 TFLIOLERRKGVQSSGIMLTFMLVALVCALALIRSKIMTALKEDAQVDLFRDITFYVYS 180

Db 121 TFLQLERKGVQSSGIMLTFMVLVYCALAIIRSKINTALKEBAQVDFPDIIFYVYFS 180
Qy 181 LLLIQLVLSCTSDSPLEFSETIHDPNCPRESSASLSTITWMTGLIVRGYRQPLESSD 240
Db 181 LLLIQLVLSCTSDSPLEFSETIHDPNCPRESSASLSTITWMTGLIVRGYRQPLESSD 240
Qy 241 LMSLNKEDTSEQVVPVLVKNMKKBCAKTRKQPVKVVYSSKDPAPQKSSSKVDANEVVAL 300
Db 241 LMSLNKEDTSEQVVPVLVKNMKKBCAKTRKQPVKVVYSSKDPAPQKSSSKVDANEVVAL 300
Qy 301 IVKSPQKMNPSLFVLYKTEGPYFLMSFPFKAIHDLMPFSGPOLIKLIIFFVNDTKAPD 360
Db 301 IVKSPQKMNPSLFVLYKTEGPYFLMSFPFKAIHDLMPFSGPOLIKLIIFFVNDTKAPD 360
Qy 361 WQGFYFVYLLFVYACLOTLVHOYFHCIFVSGMRITKAVIYAVYRKALVITNSARKSSTV 420
Db 361 WQGFYFVYLLFVYACLOTLVHOYFHCIFVSGMRITKAVIYAVYRKALVITNSARKSSTV 420
Qy 421 GEIYNLMSVDAQRFMIDLATYINMIMSAFLOYIALYILMLNGSPSVLAGVAMVLMVYVN 480
Db 421 GEIYNLMSVDAQRFMIDLATYINMIMSAFLOYIALYILMLNGSPSVLAGVAMVLMVYVN 480
Qy 481 AVMAKTKTYOVAHMKSKDNRIKLMNBILNGIKVYLKYAMELAFKDKVLAIROBELKVLK 540
Db 481 AVMAKTKTYOVAHMKSKDNRIKLMNBILNGIKVYLKYAMELAFKDKVLAIROBELKVLK 540
Qy 541 KSAVLASVGTFTWCTPPLVALCTPAAVYVITIDENNILDAQTAFAVSLAFNLRPLNLP 600
Db 541 KSAVLASVGTFTWCTPPLVALCTPAAVYVITIDENNILDAQTAFAVSLAFNLRPLNLP 600
Qy 601 MVISITVOASVSLRLRIFLSHEELEPDSIRRRPVKGGGNTSTVRNATTTMARSDPT 660
Db 601 MVISITVOASVSLRLRIFLSHEELEPDSIRRRPVKGGGNTSTVRNATTTMARSDPT 660
Qy 661 LMGITFSPREGALVAVVQVCGCGSKSLSLAEMDKVGHVAKGSVAYVPOQAMIOND 720
Db 661 LMGITFSPREGALVAVVQVCGCGSKSLSLAEMDKVGHVAKGSVAYVPOQAMIOND 720
Qy 721 SLRENILFGCQLEBPYRYSVIOACALLPDLIELPSGDTEIGEKVNLSGGQXORSILAR 780
Db 721 SLRENILFGCQLEBPYRYSVIOACALLPDLIELPSGDTEIGEKVNLSGGQXORSILAR 780
Qy 781 AVYENADIYLFDDPLSAVDAVGHIFENVIKPKMKLKNKRIILVTHSMSTLPQVDYIIV 840
Db 781 AVYENADIYLFDDPLSAVDAVGHIFENVIKPKMKLKNKRIILVTHSMSTLPQVDYIIV 840
Qy 841 MSGGKISEMGSYQELLARDGAFELRTYASTBOQDAEENGVTGVSQPKGAQOMENG 900
Db 841 MSGGKISEMGSYQELLARDGAFELRTYASTBOQDAEENGVTGVSQPKGAQOMENG 900
Qy 901 LVTDSACKOLOROLSSSSSYSGDISRRHNSTAELOKAEKKEETMKLEADKAQTGOVKL 960
Db 901 LVTDSACKOLOROLSSSSSYSGDISRRHNSTAELOKAEKKEETMKLEADKAQTGOVKL 960
Qy 961 SVYWDYKAKGLFISFISIFLPMCNHVSALASNYWLSLWDDPVLVNGTOEHTKRLSVYG 1020
Db 961 SVYWDYKAKGLFISFISIFLPMCNHVSALASNYWLSLWDDPVLVNGTOEHTKRLSVYG 1020
Qy 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLISILRSPMSFEPTPSGULVNRFSKEL 1080
Db 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLISILRSPMSFEPTPSGULVNRFSKEL 1080
Qy 1081 DTVDMSIPEVIKMFMSGLFNVIACIIVLATPIAIIIPPLGIYFFVQRFYVASSROL 1140
Db 1081 DTVDMSIPEVIKMFMSGLFNVIACIIVLATPIAIIIPPLGIYFFVQRFYVASSROL 1140
Qy 1141 KRLESVRSPIYSHFNFTLLGVSVIRAFEEBERFIHOSDLKVDENOKAYVPSIYANRWLA 1200
Db 1141 KRLESVRSPIYSHFNFTLLGVSVIRAFEEBERFIHOSDLKVDENOKAYVPSIYANRWLA 1200
Qy 1201 VRLECVCNCTIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYANMLVRRMSEMETNIVA 1260
Db 1201 VRLECVCNCTIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYANMLVRRMSEMETNIVA 1260

Db 1201 VRLECVCNCTIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYANMLVRRMSEMETNIVA 1260
Qy 1261 VERIKESYETKEKPMQIOETAPSSSWPQVGRVFPYNTCLARYEDDLPVLBHINTVING 1320
Db 1261 VERIKESYETKEKPMQIOETAPSSSWPQVGRVFPYNTCLARYEDDLPVLBHINTVING 1320
Qy 1321 EKVIGVBTGAGKSLTGLFRINESABGEIIDIIGINAKIGLHDLRKITITIPQDPVLF 1380
Db 1321 EKVIGVBTGAGKSLTGLFRINESABGEIIDIIGINAKIGLHDLRKITITIPQDPVLF 1380
Qy 1381 SGSLRMLNDPFSQYSDSEVWTSLEIAHLKDFVSLPDKLDHECABGGENLSVGQOLVCL 1440
Db 1381 SGSLRMLNDPFSQYSDSEVWTSLEIAHLKDFVSLPDKLDHECABGGENLSVGQOLVCL 1440
Qy 1441 ARALLRTKILVDEBATAVDLEFDLLOSTIRFOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Db 1441 ARALLRTKILVDEBATAVDLEFDLLOSTIRFOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Qy 1501 DKGEIOBYGAPSDLLQORGLFYSAKADAGLVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
Db 1501 DKGEIOBYGAPSDLLQORGLFYSAKADAGLVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
Qy 1561 PRYTRRRRARFVSKKNCNVAHKNIREQGRFLQDVFTTIVDLKMPHTLIIFTMSFLCSW 1620
Db 1561 PRYTRRRRARFVSKKNCNVAHKNIREQGRFLQDVFTTIVDLKMPHTLIIFTMSFLCSW 1620
Qy 1621 LLFPMVWMLIAFAHGDIAPEGNTVPCVTSIHSFSSAPLSIEVQVITIGFGRNATECP 1680
Db 1621 LLFPMVWMLIAFAHGDIAPEGNTVPCVTSIHSFSSAPLSIEVQVITIGFGRNATECP 1680
Qy 1681 LAILILIVONIVGINAINAMGCIEMKTAQAHRAEFLIISKHAVITLRHGRICPMIRVG 1740
Db 1681 LAILILIVONIVGINAINAMGCIEMKTAQAHRAEFLIISKHAVITLRHGRICPMIRVG 1740
Qy 1741 DLKSMIISATIHQVVRKTTSPGEVYVPLHQVDIPNENGVGNGIFLVAPLIYHYIDS 1800
Db 1741 DLKSMIISATIHQVVRKTTSPGEVYVPLHQVDIPNENGVGNGIFLVAPLIYHYIDS 1800
Qy 1801 NSPLYDLAPSDLLHHQOLEIIVILGEVETGTITTOARTSYLADEILMGOFVBIABED 1860
Db 1801 NSPLYDLAPSDLLHHQOLEIIVILGEVETGTITTOARTSYLADEILMGOFVBIABED 1860
Qy 1861 GRYSVDYSKFGNTIKVPTPLCTARQDLDEDSLLDALTLASSRGPILRAASVAVAKKXFS 1920
Db 1861 GRYSVDYSKFGNTIKVPTPLCTARQDLDEDSLLDALTLASSRGPILRAASVAVAKKXFS 1920
Qy 1921 ISPDLS 1927
Db 1921 ISPDLS 1927

RESULT 3
US-10-665-283-8
: Sequence 8, Application US/10665283
: Publication No. US20050063989A1
: GENERAL INFORMATION:
: APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
: APPLICANT: DERAND, Renaud
: APPLICANT: GARCIA, Elisabeth
: APPLICANT: PROST, Anne-lise
: APPLICANT: REVILLIOD, Jean
: APPLICANT: VIVAUDOU, Michel
: TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF
: FILE REFERENCE: P263US104
: CURRENT APPLICATION NUMBER: US/10/665,283
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 8
: LENGTH: 1947
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-665-283-8

Query Match 99.7%; Score 9873; DB 5; Length 1947;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1927; Conservative 0; Mismatches 0; Indels 20; Gaps 2;

```
Qy 1 MALGFCSDSDSLMDNMNTWNTSNPDFTKCFQNTVLMVPCFTLMACFFELYLSRH 60
Db 1 MALGFCSDSDSLMDNMNTWNTSNPDFTKCFQNTVLMVPCFTLMACFFELYLSRH 60
Qy 61 DRGIQMTPLNKTALGFLIMIVCMADLFYSFMRSGIFLAPVFLVSPILLGITTLA 120
Db 61 DRGIQMTPLNKTALGFLIMIVCMADLFYSFMRSGIFLAPVFLVSPILLGITTLA 120
Qy 61 DRGIQMTPLNKTALGFLIMIVCMADLFYSFMRSGIFLAPVFLVSPILLGITTLA 120
Db 61 DRGIQMTPLNKTALGFLIMIVCMADLFYSFMRSGIFLAPVFLVSPILLGITTLA 120
Qy 121 TFLIQLEBRKGVSSGIMLTFMLVALYCALILRSKIMTALKEDAQVDLPFDITFYVFS 180
Db 121 TFLIQLEBRKGVSSGIMLTFMLVALYCALILRSKIMTALKEDAQVDLPFDITFYVFS 180
Qy 181 LLLIQVLVSCFSDSPLEFSETIHDNPPCRESSASLSTITFWMTGLIVRGYRQPLESD 240
Db 181 LLLIQVLVSCFSDSPLEFSETIHDNPPCRESSASLSTITFWMTGLIVRGYRQPLESD 240
Qy 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVVAL 300
Db 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVVAL 300
Qy 301 IVKSPQKEMNDSLFRVLYKTFGPYFLMSFFFKAIHDLMMFSGPOILLKILKRVNDTKAPD 360
Db 301 IVKSPQKEMNDSLFRVLYKTFGPYFLMSFFFKAIHDLMMFSGPOILLKILKRVNDTKAPD 360
Qy 361 WQGFYVYVLLFVTAQLOVLVHOYHICFVSGMRKKTAVIGAVYRKALVITNSARKSTV 420
Db 361 WQGFYVYVLLFVTAQLOVLVHOYHICFVSGMRKKTAVIGAVYRKALVITNSARKSTV 420
Qy 421 GEIYNLMSVDQRFMDLATYINMISAPLOYILALYLMNLGSPVLGAVVMVLVNVN 480
Db 421 GEIYNLMSVDQRFMDLATYINMISAPLOYILALYLMNLGSPVLGAVVMVLVNVN 480
Qy 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAMBLAFKQVLAIROBELKYLK 540
Db 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAMBLAFKQVLAIROBELKYLK 540
Qy 541 KSAVLASANGFTWCTPFLVALCTPAVVYVITDENNIIDAQTAFLNLIRPLNLTP 600
Db 541 KSAVLASANGFTWCTPFLVALCTPAVVYVITDENNIIDAQTAFLNLIRPLNLTP 600
Qy 601 MVISSIVQASVSLKRLRIFLSHEBLEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHEBLEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
Qy 661 LMGITFSP1PBGALVAVVQVCGKSSLLSALLAEMDKVEGHVAKGSVAVYVPOQAMIOND 720
Db 661 LMGITFSP1PBGALVAVVQVCGKSSLLSALLAEMDKVEGHVAKGSVAVYVPOQAMIOND 720
Qy 721 SLRENILFQCQLEBEYHRSVIOACALLPDLBILBEGDRTIEGKCVNLSGGQKORVSLAR 780
Db 721 SLRENILFQCQLEBEYHRSVIOACALLPDLBILBEGDRTIEGKCVNLSGGQKORVSLAR 780
Qy 781 AVYSNADLYLFDDPLSAVDAHVKHIFENVIKPKMLKNKTRIIIVTHSMVLPQVDVITV 840
Db 781 AVYSNADLYLFDDPLSAVDAHVKHIFENVIKPKMLKNKTRIIIVTHSMVLPQVDVITV 840
Qy 841 MSGGKISEMGSYOEILARDGAFAEFLRTYASTEQDADAEENGVTGVSQPGKEAKOMENGM 900
Db 841 MSGGKISEMGSYOEILARDGAFAEFLRTYASTEQDADAEENGVTGVSQPGKEAKOMENGM 900
Qy 901 LVTTDSAGQOLQROUSSSSSYSGDISRHHNSTAELOKAPAKKEETWKLMEADKAOQGYVL 960
Db 901 LVTTDSAGQOLQROUSSSSSYSGDISRHHNSTAELOKAPAKKEETWKLMEADKAOQGYVL 960
Qy 961 SVYMDYMAKIGLFTSFLSIFLPMCHVASLASNYWLSLMTDPIVNGTQEHKTVLSYVG 1020
Db 961 SVYMDYMAKIGLFTSFLSIFLPMCHVASLASNYWLSLMTDPIVNGTQEHKTVLSYVG 1020
```

```
Qy 1021 ALGISQIANVFGYSMANVSGIILASRCLHYDLSHIIIRSPMSFFERTPSGULVNRFSKEL 1080
Db 1021 ALGISQIANVFGYSMAVSGIILASRCLHYDLSHIIIRSPMSFFERTPSGULVNRFSKEL 1080
Qy 1081 DTVDMSIPEVIKMGMSLFNVIGACIVILLATPAAIILIPPLGIFYFVORFYAASSROL 1140
Db 1081 DTVDMSIPEVIKMGMSLFNVIGACIVILLATPAAIILIPPLGIFYFVORFYAASSROL 1140
Qy 1141 KRLESVSRSPVSHFNETLLGVSVIRAFEEQEREIHOSSDKVDENQAKAYPSIVANRWLA 1200
Db 1141 KRLESVSRSPVSHFNETLLGVSVIRAFEEQEREIHOSSDKVDENQAKAYPSIVANRWLA 1200
Qy 1201 VRLBCVNCIYLPALPAVISRHSLSAGLVLSYSLSQVTTYNMLVRNSEMETNIVA 1260
Db 1201 VRLBCVNCIYLPALPAVISRHSLSAGLVLSYSLSQVTTYNMLVRNSEMETNIVA 1260
Qy 1261 VERLKEYSETEKEAPWQIOETAPSSMPQVRVEFRNYCLRYRBDLPVLRHIVNTINGG 1320
Db 1261 VERLKEYSETEKEAPWQIOETAPSSMPQVRVEFRNYCLRYRBDLPVLRHIVNTINGG 1320
Qy 1321 EKVGIVGRGTGAGKSLTLGLFRINESAGEIILIDGINIAKIGLHDLRFKTIILIPQDPYLF 1380
Db 1321 EKVGIVGRGTGAGKSLTLGLFRINESAGEIILIDGINIAKIGLHDLRFKTIILIPQDPYLF 1380
Qy 1381 SGSLRNULDPPSQYSDEEWTSLELAHKDFVSALPKLDBHECABGEGNLSVGGROLYCL 1440
Db 1381 SGSLRNULDPPSQYSDEEWTSLELAHKDFVSALPKLDBHECABGEGNLSVGGROLYCL 1440
Qy 1441 ABALLRKTILVDEAPAAVDELTDDLIQSTIRQFEDCTVLTAAHLNLTMDTRIVYL 1500
Db 1441 ABALLRKTILVDEAPAAVDELTDDLIQSTIRQFEDCTVLTAAHLNLTMDTRIVYL 1500
Qy 1501 DKGEIOEYGAPSDLLQORGLFYSMAXDAGLVGGGGGMLSRKGIIPREYVLTIRLAEDPAE 1560
Db 1501 DKGEIOEYGAPSDLLQORGLFYSMAXDAGLVGGGGGMLSRKGIIPREYVLTIRLAEDPAE 1560
Qy 1561 PRYTRERRARFVSKKGCNVAKHNIIEQGRFLQDVFTTVLDLKWPHULLFTMSFLCSW 1620
Db 1561 PRYTRERRARFVSKKGCNVAKHNIIEQGRFLQDVFTTVLDLKWPHULLFTMSFLCSW 1620
Qy 1621 LLFPMWMLTAPAG-----DLAP-----GBGNVPCVTSIHSFSAFLF 1660
Db 1621 LLFPMWMLTAPAG-----DLAP-----GBGNVPCVTSIHSFSAFLF 1660
Qy 1661 SIEVOVITGFGGRNATECPALILILIVONIVGLMINAMLCIFMNTAQAHRAETLIF 1720
Db 1661 SIEVOVITGFGGRNATECPALILILIVONIVGLMINAMLCIFMNTAQAHRAETLIF 1720
Qy 1721 SKHAVITLRHGRLCFPMLRVGDLRKSMIISATIMQVVRKTTSPGGEVPLHQVDIPMEHG 1780
Db 1721 SKHAVITLRHGRLCFPMLRVGDLRKSMIISATIMQVVRKTTSPGGEVPLHQVDIPMEHG 1780
Qy 1781 VGGNGIFLAPLIIYHYIDNSPLYDLAPSLDHHODLEIIVILEGVEITGITTOARTS 1840
Db 1781 VGGNGIFLAPLIIYHYIDNSPLYDLAPSLDHHODLEIIVILEGVEITGITTOARTS 1840
Qy 1841 YLADEIILMGQRFVPIVAEEDGRYSVDYSKFENITIKVPTPLCTAQLDDEBSLDAITLAS 1900
Db 1841 YLADEIILMGQRFVPIVAEEDGRYSVDYSKFENITIKVPTPLCTAQLDDEBSLDAITLAS 1900
Qy 1901 SRGPLRRRSVAVAKKPFSSIPDLS 1927
Db 1901 SRGPLRRRSVAVAKKPFSSIPDLS 1927
```

RESULT 4
US-10-665-283-4
; Sequence 4, Application US/10665283
; Publication No. US2005063989A1
; GENERAL INFORMATION:
; APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
; APPLICANT: DERAND, Renaud
; APPLICANT: GARCIA, Elisabeth

; APPLICANT: PROST, Anne-Lise
 ; APPLICANT: REVILLOU, Jean
 ; APPLICANT: VIVAUDOU, Michel
 ; TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF
 ; FILE REFERENCE: F26305104
 ; CURRENT APPLICATION NUMBER: US/10/665,283
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 1891
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-665-283-4

Query Match 98.3%; Score 9734; DB 5; Length 1891;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALGFCGADSDPLMDNMTNMTNSPDTCFQNTVYVWPCFYLACFPPEYLYSRH 60
 DB 1 MALGFCGADSDPLMDNMTNMTNSPDTCFQNTVYVWPCFYLACFPPEYLYSRH 60
 QY 61 DRGYIOMTPLNKTALGFLMIYCMADLFYSFMRSGIFLAPYFLVSPITLLGTTLLA 120
 DB 61 DRGYIOMTPLNKTALGFLMIYCMADLFYSFMRSGIFLAPYFLVSPITLLGTTLLA 120
 QY 121 TPLIQLERRKQVSSGIMLTFWLVALCALILRSKIMTALKEDAQVDLPFDITFYVFS 180
 DB 121 TPLIQLERRKQVSSGIMLTFWLVALCALILRSKIMTALKEDAQVDLPFDITFYVFS 180
 QY 181 LLLIOLVLSGDSPLSETHIDNPNCPSSASLSTITFWMTGLIVRGYRQPLESSD 240
 DB 181 LLLIOLVLSGDSPLSETHIDNPNCPSSASLSTITFWMTGLIVRGYRQPLESSD 240
 QY 241 LMSLNKEDTSEQVAVLVKNNKKECAKTRKQPVKVVYSSKDPAPKSSKVDANEVEAL 300
 DB 241 LMSLNKEDTSEQVAVLVKNNKKECAKTRKQPVKVVYSSKDPAPKSSKVDANEVEAL 300
 QY 301 IVKSPQKEMNDSEFKVLYKTGPFYFLMSFFKAIHDLMMFSGPOLIKLIRFVNDTKAD 360
 DB 301 IVKSPQKEMNDSEFKVLYKTGPFYFLMSFFKAIHDLMMFSGPOLIKLIRFVNDTKAD 360
 QY 361 WQGFYVYLLFETACLOTLVHOYTHICFVSGMRKKTAVIGAVRKALVITNSARKSTV 420
 DB 361 WQGFYVYLLFETACLOTLVHOYTHICFVSGMRKKTAVIGAVRKALVITNSARKSTV 420
 QY 421 GEIYVLSVDAQRFMDLATYINMIWSAPLOYILALYLLMLNIGPSVLGAVVMTLMEVYN 480
 DB 421 GEIYVLSVDAQRFMDLATYINMIWSAPLOYILALYLLMLNIGPSVLGAVVMTLMEVYN 480
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKAYAMELAFKDKYALIROBELKYLK 540
 DB 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKAYAMELAFKDKYALIROBELKYLK 540
 QY 541 KSAIYLSAVGTCTTWCPTPLVALCTPAVYVITIDENNILDAQAFYSALFNILRPPLNLP 600
 DB 541 KSAIYLSAVGTCTTWCPTPLVALCTPAVYVITIDENNILDAQAFYSALFNILRPPLNLP 600
 QY 601 MVISSIVOASVSLKRLRIFLSHEELEPDSIRRPVKDGGGNSITVRNATVTMARSDPT 660
 DB 601 MVISSIVOASVSLKRLRIFLSHEELEPDSIRRPVKDGGGNSITVRNATVTMARSDPT 660
 QY 661 LNTGTFSTPEGALVAVVGQVCGKSSLSALLAEMDKVGHVAIKGSVAVYPOQAWIQND 720
 DB 661 LNTGTFSTPEGALVAVVGQVCGKSSLSALLAEMDKVGHVAIKGSVAVYPOQAWIQND 720
 QY 721 SLRNNILFGCOLLEPYRVSVOACALLPDLIELSGDTEIGEGVNLSGGQKQVSLAR 780
 DB 721 SLRNNILFGCOLLEPYRVSVOACALLPDLIELSGDTEIGEGVNLSGGQKQVSLAR 780
 QY 781 AVYSNADILYLPDDPLSAVDAHVGHKIFENVIGPKMLKNKTRILVTHSMSTLPQVDVILV 840

DB 781 AVYSNADILYLPDDPLSAVDAHVGHKIFENVIGPKMLKNKTRILVTHSMSTLPQVDVILV 840
 QY 841 MSGGKISMSGYQOILLARDGFAFFLFTYATBEOBDAENGVCVSGPGKAAQOMENG 900
 DB 841 MSGGKISMSGYQOILLARDGFAFFLFTYATBEOBDAENGVCVSGPGKAAQOMENG 900
 QY 901 LVYDSAGKOLQROUSSSSSGYSDISRHNSTAELOKAEAKKEETWKMEAKAQTQYKL 960
 DB 901 LVYDSAGKOLQROUSSSSSGYSDISRHNSTAELOKAEAKKEETWKMEAKAQTQYKL 960
 QY 961 SVYWDYKAIGLFISFISIFLMCNHVSALASNWLSLMTDPIVNGTOBHTKYL 1020
 DB 961 SVYWDYKAIGLFISFISIFLMCNHVSALASNWLSLMTDPIVNGTOBHTKYL 1020
 QY 1021 ALGISOGIAYVGYMAVSIIGITLASRCLHVDLHSLIRSPMSFFERTPSGULVNRFSKEL 1080
 DB 1021 ALGISOGIAYVGYMAVSIIGITLASRCLHVDLHSLIRSPMSFFERTPSGULVNRFSKEL 1080
 QY 1081 DTVDKMTPEVTKMFMGSLFENVIGACIVITLATAPIAIIIPPLGLIFPVORFYVASSROL 1140
 DB 1081 DTVDKMTPEVTKMFMGSLFENVIGACIVITLATAPIAIIIPPLGLIFPVORFYVASSROL 1140
 QY 1141 KRLSVSRSPYSHFNETLGVSVYIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
 DB 1141 KRLSVSRSPYSHFNETLGVSVYIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
 QY 1201 VRLCEVGNCTVLPALPAVYISRHSLAGVLGSLYSLOVTTYLNMLVRMSEMETNIVA 1260
 DB 1201 VRLCEVGNCTVLPALPAVYISRHSLAGVLGSLYSLOVTTYLNMLVRMSEMETNIVA 1260
 QY 1261 VERLKEYSETEKAPMOIQETAPSSWPQVGRVFRNYCLARYEDLPVLRPHINTNGG 1320
 DB 1261 VERLKEYSETEKAPMOIQETAPSSWPQVGRVFRNYCLARYEDLPVLRPHINTNGG 1320
 QY 1321 EKVGIQVGTGAGKSLITLGLFRINSAEGEIIIDIGINIAKGLHDLRFKTIIPQDPVLF 1380
 DB 1321 EKVGIQVGTGAGKSLITLGLFRINSAEGEIIIDIGINIAKGLHDLRFKTIIPQDPVLF 1380
 QY 1381 SGSLRNMULDPSQYSDSEWMTSLTALHKDFVSALPKLDHECAEGENLSVGQROLVCL 1440
 DB 1381 SGSLRNMULDPSQYSDSEWMTSLTALHKDFVSALPKLDHECAEGENLSVGQROLVCL 1440
 QY 1441 ARALLRKTILVDEATAVDLETDLIQSTIRQFEDCTVLTIAHRLNTIMDTRYIVL 1500
 DB 1441 ARALLRKTILVDEATAVDLETDLIQSTIRQFEDCTVLTIAHRLNTIMDTRYIVL 1500
 QY 1501 DKGEIQRYGAPSDLLQORGLFYSAKXAGLVGGGGGMLSRKGIIPBEYVUTRIAEDEPAE 1560
 DB 1501 DKGEIQRYGAPSDLLQORGLFYSAKXAGLVGGGGGMLSRKGIIPBEYVUTRIAEDEPAE 1560
 QY 1561 PRYTRERRARFVSKKNCNVAAHKNIREQGRFLQDVFTTVLDLWPHULLIPTMSFLCSW 1620
 DB 1561 PRYTRERRARFVSKKNCNVAAHKNIREQGRFLQDVFTTVLDLWPHULLIPTMSFLCSW 1620
 QY 1621 LIFAMVWMLIAFAGDILAPBGTNVPCVTSIHSFSAFLPSIEQVITIGFGRWTEBEP 1680
 DB 1621 LIFAMVWMLIAFAGDILAPBGTNVPCVTSIHSFSAFLPSIEQVITIGFGRWTEBEP 1680
 QY 1681 LAIILILVQNIQVGMIAVIMGCIIPKTAQAHRAETLIFSIAVITLIRHGRCLCFMLRVG 1740
 DB 1681 LAIILILVQNIQVGMIAVIMGCIIPKTAQAHRAETLIFSIAVITLIRHGRCLCFMLRVG 1740
 QY 1741 DLKRSMTISATIHQVVRKTTSPGEVAVPLHQVDIPMENGVGNGIFLVAPLIYHVIDS 1800
 DB 1741 DLKRSMTISATIHQVVRKTTSPGEVAVPLHQVDIPMENGVGNGIFLVAPLIYHVIDS 1800
 QY 1801 NSPLYDLAPSDLHHQDLIELIVILEGVENTGTITQARTSLADEILMGQRFVVAEED 1860
 DB 1801 NSPLYDLAPSDLHHQDLIELIVILEGVENTGTITQARTSLADEILMGQRFVVAEED 1860
 QY 1861 GRYSVDYSKFGNTIKVPTPLCTARQLDEDRS 1891
 DB 1861 GRYSVDYSKFGNTIKVPTPLCTARQLDEDRS 1891

RESULT 5
US-09-939-853A-86
; Sequence 86, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-86

Query Match 79.4%; Score 7860; DB 3; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCADSGSDPLMDNMVNTMTNSPDTKCFONTVLVWVCFYLMACFPEFYLYLSRH 60
DB 1 MALRGFCADSGSDPLMDNMVNTMTNSPDTKCFONTVLVWVCFYLMACFPEFYLYLSRH 60
QY 61 DRGIQMTPLNKTATGELLMTVCADLFYSFWERSGFIPLAVFLVSPITLGIITLLA 120
DB 61 DRGIQMTPLNKTATGELLMTVCADLFYSFWERSGFIPLAVFLVSPITLGIITLLA 120
QY 121 TELLQERKGVQSSGIMLTFWLVAVCALATLRKIMTALKEQOVDFRDTYYVFS 180
DB 121 TELLQERKGVQSSGIMLTFWLVAVCALATLRKIMTALKEQOVDFRDTYYVFS 180
QY 181 LLLIQLVSCFSDRSLFSETHDNPCESSASFLSRITTFWITGLIVRGYRQPLEGSD 240
DB 181 LLLIQLVSCFSDRSLFSETHDNPCESSASFLSRITTFWITGLIVRGYRQPLEGSD 240
QY 241 LMSLINKEDTSQVVPVLVKNMKKECAKTRKQPVKVVYSSKDBAQPKSSKVDANEVEAL 300
DB 241 LMSLINKEDTSQVVPVLVKNMKKECAKTRKQPVKVVYSSKDBAQPKSSKVDANEVEAL 300
QY 301 IVKSPQKMNPSLFFVLYKTGPFLMSPFKAHIDLMFGSPQILKLIKRVNDTKAPD 360
DB 301 IVKSPQKMNPSLFFVLYKTGPFLMSPFKAHIDLMFGSPQILKLIKRVNDTKAPD 360
QY 361 WOGYYTYLLFVTACLOTLVHQQYHICFVSGMRKTAIVIGAVYKALVITNSAKSSIV 420
DB 361 WOGYYTYLLFVTACLOTLVHQQYHICFVSGMRKTAIVIGAVYKALVITNSAKSSIV 420
QY 421 GEIVNLSVDAQRFMDLATYINMISAPLOVILALYLLMLNGPSVLGAVAVMLVAVN 480
DB 421 GEIVNLSVDAQRFMDLATYINMISAPLOVILALYLLMLNGPSVLGAVAVMLVAVN 480
QY 481 AVMAKTKTYOYAHMKSNDRIKLMNEIINGIKVLTAYAMELAFQDKVLATROEELKYLK 540
DB 481 AVMAKTKTYOYAHMKSNDRIKLMNEIINGIKVLTAYAMELAFQDKVLATROEELKYLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCFEAVYVTDENNIIILAAQTAFFVSLALFNIRPPLNIP 600
DB 541 KSAVLSAVGTFTWCTPFLVALCFEAVYVTDENNIIILAAQTAFFVSLALFNIRPPLNIP 600
QY 601 MVISSIVQASVSLKRLIFLSHELEPPDSIERRPVKDGGGTNSITVRNATFTWASDDEPT 660

DB 601 MVISSIVQASVSLKRLIFLSHELEPPDSIERRPVKDGGGTNSITVRNATFTWASDDEPT 660
QY 661 LMGITF6IPEGALVAVVGVCGCKSSLSLALAEADKVEGVAIKGSVAVYPQOAMIOND 720
DB 661 LMGITF6IPEGALVAVVGVCGCKSSLSLALAEADKVEGVAIKGSVAVYPQOAMIOND 720
QY 721 SLRENILFGCQLEBPYRSVIOACALLPDLIELPSGRTEIGERGVNLSGGQKORVSLAR 780
DB 721 SLRENILFGCQLEBPYRSVIOACALLPDLIELPSGRTEIGERGVNLSGGQKORVSLAR 780
QY 781 AVYNNADIIYLFDDPLSAVDVAVGHIENYVIGPKMLKNKRIILVTHSMSTLPQVDVIV 840
DB 781 AVYNNADIIYLFDDPLSAVDVAVGHIENYVIGPKMLKNKRIILVTHSMSTLPQVDVIV 840
QY 841 MSGGKISEMGSYQELIARDGAFAEFLRTYASTQEQODAEENGTVGSGPGKEAKOMENG 900
DB 841 MSGGKISEMGSYQELIARDGAFAEFLRTYASTQEQODAEENGTVGSGPGKEAKOMENG 900
QY 901 LVYDSAGKOLQROLSSSSSSYSGDISRHNSYAEIQAAPAKKEFTWKLMEADKAQTGVYL 960
DB 901 LVYDSAGKOLQROLSSSSSSYSGDISRHNSYAEIQAAPAKKEFTWKLMEADKAQTGVYL 960
QY 961 SVYWDYMKALGLFISFLSIFLMCNHVSALASNYWLSLMTDPIVNGTQEHTKVRLSYG 1020
DB 961 SVYWDYMKALGLFISFLSIFLMCNHVSALASNYWLSLMTDPIVNGTQEHTKVRLSYG 1020
QY 1021 ALGISOGIAVFGYMAVSIIGIILASRCIHDVLIHSILRSPMSPFERPNSGMLVNRFSKEL 1080
DB 1021 ALGISOGIAVFGYMAVSIIGIILASRCIHDVLIHSILRSPMSPFERPNSGMLVNRFSKEL 1080
QY 1081 DTVDMSIPEVIKMGMSLFENVIGACIVILLATPIAIIIPPLGIYFFVQGFYVASSRQL 1140
DB 1081 DTVDMSIPEVIKMGMSLFENVIGACIVILLATPIAIIIPPLGIYFFVQGFYVASSRQL 1140
QY 1141 KRLSSVSSPYVSHPNETLLGVSTIRAFEEBERITHOSDLKVDENOKAYVPSIVANRWLA 1200
DB 1141 KRLSSVSSPYVSHPNETLLGVSTIRAFEEBERITHOSDLKVDENOKAYVPSIVANRWLA 1200
QY 1201 VRLECVGNCIYLFALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRRSMSEMETNIVA 1260
DB 1201 VRLECVGNCIYLFALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRRSMSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPPSWPQVGRVFRNYCLARYEDLDVLAHINVTINGG 1320
DB 1261 VERLKEYSETEKEAPWQIOETAPPSWPQVGRVFRNYCLARYEDLDVLAHINVTINGG 1320
QY 1321 EKVGIYGRGTGAKSSLTGLFRINESAGEIIDIIGINIAKGLHDLRFXKIIIPQDPVLF 1380
DB 1321 EKVGIYGRGTGAKSSLTGLFRINESAGEIIDIIGINIAKGLHDLRFXKIIIPQDPVLF 1380
QY 1381 SSSLRMLNDPFSQVSDERVWTSLELAHLKDVSAALPDLDBECAGGENLSVGOQOLVCL 1440
DB 1381 SSSLRMLNDPFSQVSDERVWTSLELAHLKDVSAALPDLDBECAGGENLSVGOQOLVCL 1440
QY 1441 ARALLRKTILVLDEATAVADLETDLLOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
DB 1441 ARALLRKTILVLDEATAVADLETDLLOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
QY 1501 DKGEIOEYGAQSDLIQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGAQSDLIQORGLFYSMAKDAGLV 1531

RESULT 6
US-10-667-891-6
; Sequence 6, Application US/10667891
; Publication No. US20040171024A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, CHARLES W.
; APPLICANT: BREY, PAUL T.
; APPLICANT: HOLM, INGÉ
; APPLICANT: GRAILLES, MARINE

; APPLICANT: RZHEITSKY, ANDREY
 ; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND
 ; FILE OF INVENTION: ANOPHELES
 ; FILE REFERENCE: 03495-0294-00000
 ; CURRENT APPLICATION NUMBER: US/10/667,891
 ; CURRENT FILING DATE: 2003-09-23
 ; PRIOR APPLICATION NUMBER: 60/413,469
 ; PRIOR FILING DATE: 2002-09-26
 ; NUMBER OF SEQ ID NOS: 76
 ; SOFTWARE: PatentIn Ver. 3.2
 ; SEQ ID NO: 6
 ; LENGTH: 1531
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-667-891-6

Query Match 79.4%; Score 7860; DB 4; Length 1531;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFSGADSDPLMDMNTMTNTSNDPTKCFQNTYLVWPCGYLMACFPFYLYSRH 60
 DB 1 MALRGFSGADSDPLMDMNTMTNTSNDPTKCFQNTYLVWPCGYLMACFPFYLYSRH 60
 QY 61 DRGYIOMTPINKTKTALGFLMIYCMADLFYSFWEBSRGIFLAVPLVSPFLIGITTLA 120
 DB 61 DRGYIOMTPINKTKTALGFLMIYCMADLFYSFWEBSRGIFLAVPLVSPFLIGITTLA 120
 QY 121 TFLIOLERRKGVQSSGIMLTFMVALVCAALILRSKIMTALKEDAQVDLPFDITFYVFS 180
 DB 121 TFLIOLERRKGVQSSGIMLTFMVALVCAALILRSKIMTALKEDAQVDLPFDITFYVFS 180
 QY 181 LLLIQLVLSGFSDBSPFSETIHDNPCESSASFLSRITFTWTGILVRCYRPLEGSD 240
 DB 181 LLLIQLVLSGFSDBSPFSETIHDNPCESSASFLSRITFTWTGILVRCYRPLEGSD 240
 QY 181 LLLIQLVLSGFSDBSPFSETIHDNPCESSASFLSRITFTWTGILVRCYRPLEGSD 240
 DB 181 LLLIQLVLSGFSDBSPFSETIHDNPCESSASFLSRITFTWTGILVRCYRPLEGSD 240
 QY 241 LMSLNKEDTSEOVVPLVYKWKKECAKTRKOPVYVYSSKDPAPKESKVDANEVEAL 300
 DB 241 LMSLNKEDTSEOVVPLVYKWKKECAKTRKOPVYVYSSKDPAPKESKVDANEVEAL 300
 QY 301 IVKSPQKEMNPSPFLKVLTKTFGPYFLMSFPEKAIHDLMPGPOILKLLIKFVNDTRAPD 360
 DB 301 IVKSPQKEMNPSPFLKVLTKTFGPYFLMSFPEKAIHDLMPGPOILKLLIKFVNDTRAPD 360
 QY 361 WQGFYTVLLFTVACLOTLYLHQYFHCIFVSGMRKTAIVGAYVRKALVITNSARKSTV 420
 DB 361 WQGFYTVLLFTVACLOTLYLHQYFHCIFVSGMRKTAIVGAYVRKALVITNSARKSTV 420
 QY 421 GEIYNLMSVDAORFMDLATYINMTWSAPLOYILALYLLMLTGPSTVAGVAVVMTLMPVN 480
 DB 421 GEIYNLMSVDAORFMDLATYINMTWSAPLOYILALYLLMLTGPSTVAGVAVVMTLMPVN 480
 QY 481 AVNAAKTKTYQVAMHMSKDNRIKLMNEILNGIKVLYKLYAMELAFKDKVLAIROBELKVLK 540
 DB 481 AVNAAKTKTYQVAMHMSKDNRIKLMNEILNGIKVLYKLYAMELAFKDKVLAIROBELKVLK 540
 QY 541 KSAVLAVGFTFWCTPFLVALCTFAVYVITIDENNIIDAQTAFAVSALFNILREPLNLP 600
 DB 541 KSAVLAVGFTFWCTPFLVALCTFAVYVITIDENNIIDAQTAFAVSALFNILREPLNLP 600
 QY 601 MVTSSIVQASVSLKRLIFLSHEBLEPDSIRRPVKGGGNSITVNNATFTWARSPT 660
 DB 601 MVTSSIVQASVSLKRLIFLSHEBLEPDSIRRPVKGGGNSITVNNATFTWARSPT 660
 QY 661 LINGITFISPEGALVAVVGVCCKSSLLSALLAEMDKVEGHVAIKGSVAVVPOQAWIOND 720
 DB 661 LINGITFISPEGALVAVVGVCCKSSLLSALLAEMDKVEGHVAIKGSVAVVPOQAWIOND 720
 QY 721 SLRENIIFGCOLLEPPYRSYIACALPDLIELPSGRTEIGEGVNLGGQOKRVSLAR 780
 DB 721 SLRENIIFGCOLLEPPYRSYIACALPDLIELPSGRTEIGEGVNLGGQOKRVSLAR 780
 QY 781 AVYSNADIYLFDDPLSAVDHVGHIPEENVIGPKGMLKNKTRILLVTHSMYSGLPQVDYIIV 840

DB 781 AVYSNADIYLFDDPLSAVDHVGHIPEENVIGPKGMLKNKTRILLVTHSMYSGLPQVDYIIV 840
 QY 841 MSGGKISEMSYQELLARDGAFAPLRTYASTROBDAENGYTGSNGPEKAKOMENGM 900
 DB 841 MSGGKISEMSYQELLARDGAFAPLRTYASTROBDAENGYTGSNGPEKAKOMENGM 900
 QY 901 LVYDSAGKOLQORLSSSSSSGDIISRHNSIAELQKAEAKKEFTWKLMEADKQOTGVKL 960
 DB 901 LVYDSAGKOLQORLSSSSSSGDIISRHNSIAELQKAEAKKEFTWKLMEADKQOTGVKL 960
 QY 961 SVYWDYWKALGFLISFLSIFLPMCNHVSALASNYWLSMTDDBIUNGTQHTKRLSVYG 1020
 DB 961 SVYWDYWKALGFLISFLSIFLPMCNHVSALASNYWLSMTDDBIUNGTQHTKRLSVYG 1020
 QY 1021 ALGISOGIAVFGYSMAVISIGIILASRCLHVDLHSLIRSPMSFERTPSGNLVNRFSEKL 1080
 DB 1021 ALGISOGIAVFGYSMAVISIGIILASRCLHVDLHSLIRSPMSFERTPSGNLVNRFSEKL 1080
 QY 1081 DTVDSMIPYIKMFMGSLFVIGACIVILATPIAIIIPPLGLIYFVQRFYVASSROL 1140
 DB 1081 DTVDSMIPYIKMFMGSLFVIGACIVILATPIAIIIPPLGLIYFVQRFYVASSROL 1140
 QY 1141 KRLSEYSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
 DB 1141 KRLSEYSRSPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
 QY 1201 VRLFCVNCIIVLPAALFAVISRHSLSAGLVGSVSYLQVTTYIAMLVRMSSEMETNIVA 1260
 DB 1201 VRLFCVNCIIVLPAALFAVISRHSLSAGLVGSVSYLQVTTYIAMLVRMSSEMETNIVA 1260
 QY 1261 VERLKEYSETEKAPMOIQIOTAPSSMPQVGRVFNRYCIRYREDDLPVLRHINVTING 1320
 DB 1261 VERLKEYSETEKAPMOIQIOTAPSSMPQVGRVFNRYCIRYREDDLPVLRHINVTING 1320
 QY 1321 EKVGIQVGTAGKSSLTGLFGRINESAGEIIDIINAIKIGLHDLRFKTTIIPQDPVLF 1380
 DB 1321 EKVGIQVGTAGKSSLTGLFGRINESAGEIIDIINAIKIGLHDLRFKTTIIPQDPVLF 1380
 QY 1381 SCSLRNMLDFSOYSDEEWTSLAHLKDFVSALPDKLDHECAGEGENISVGOQLVCL 1440
 DB 1381 SCSLRNMLDFSOYSDEEWTSLAHLKDFVSALPDKLDHECAGEGENISVGOQLVCL 1440
 QY 1441 ARALLKTKILVNDERTAAVDETDLIQSTIRTOBEDCTVLTARLNTIMYTRYVL 1500
 DB 1441 ARALLKTKILVNDERTAAVDETDLIQSTIRTOBEDCTVLTARLNTIMYTRYVL 1500
 QY 1501 DKGEIOEYGAPSDLOORGFLFYSMAKDAGLV 1531
 DB 1501 DKGEIOEYGAPSDLOORGFLFYSMAKDAGLV 1531

RESULT 7
 US-10-807-466-6
 ; Sequence 6, Application US/10807466
 ; Publication No. US20040244066A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ROTH, CHARLES W.
 ; APPLICANT: BREV, PAUL T.
 ; APPLICANT: HOLM, INGE
 ; APPLICANT: GRALLIES, MARINE
 ; APPLICANT: RZHEITSKY, ANDREY
 ; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND
 ; FILE OF INVENTION: ANOPHELES
 ; FILE REFERENCE: 03495-0294-01000
 ; CURRENT APPLICATION NUMBER: US/10/807,466
 ; CURRENT FILING DATE: 2004-03-24
 ; PRIOR APPLICATION NUMBER: 10/667,891
 ; PRIOR FILING DATE: 2003-09-23
 ; PRIOR APPLICATION NUMBER: 60/413,469
 ; PRIOR FILING DATE: 2002-09-26
 ; NUMBER OF SEQ ID NOS: 139
 ; SOFTWARE: PatentIn Ver. 3.2

```
; SEQ ID NO 6
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-807-466-6

Query Match      79.4%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MALRGFCGADSDPLMDMNNVTMTNSNPDFTKCFONTVLVWVPCFYLMACFPFFFLYLSRH 60
DB      1 MALRGFCGADSDPLMDMNNVTMTNSNPDFTKCFONTVLVWVPCFYLMACFPFFFLYLSRH 60
QY      61 DRGIQWMTPLNKTATLGFLLIMVCMADLFYSFWERSRGIFLAPVFLVSPILLGITTLTA 120
DB      61 DRGIQWMTPLNKTATLGFLLIMVCMADLFYSFWERSRGIFLAPVFLVSPILLGITTLTA 120
QY      121 TFLIQLEERRKGVSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLPFDITFYVYS 180
DB      121 TFLIQLEERRKGVSSGIMLTFWLVALVCAALILRSKIMTALKEDAQVDLPFDITFYVYS 180
QY      181 LLLIQVLVSCFSDSPSEFTIHDPNCPBESSASFSLITFMWTGLIVRGVROPLBSSD 240
DB      181 LLLIQVLVSCFSDSPSEFTIHDPNCPBESSASFSLITFMWTGLIVRGVROPLBSSD 240
QY      241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVKVYSSKDPAPCKESSKYDANEVEAL 300
DB      241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVKVYSSKDPAPCKESSKYDANEVEAL 300
QY      301 IVKSPQKEMNPSLFKVLYKTGPFYLMSPFFKAIHDLMMFSGPOLIKLILFVNDTKAPD 360
DB      301 IVKSPQKEMNPSLFKVLYKTGPFYLMSPFFKAIHDLMMFSGPOLIKLILFVNDTKAPD 360
QY      361 MOGFYFYTLLFVTACLOTLVHOVPHICFVSGMRKKTAVIGAVYKALVITNSARKSTV 420
DB      361 MOGFYFYTLLFVTACLOTLVHOVPHICFVSGMRKKTAVIGAVYKALVITNSARKSTV 420
QY      421 GEIYVNLVSDAQRFMDLATYINMIWSAPLOVILALYLLMLNGPSVLGAVVMVLMVEVN 480
DB      421 GEIYVNLVSDAQRFMDLATYINMIWSAPLOVILALYLLMLNGPSVLGAVVMVLMVEVN 480
QY      481 AVMAKTKTYOVAAHKSKONRIKLMNEILNGIKVLKLIYAMELAFDXYALAIROEELKYLK 540
DB      481 AVMAKTKTYOVAAHKSKONRIKLMNEILNGIKVLKLIYAMELAFDXYALAIROEELKYLK 540
QY      541 KSAVLSAAGTFTWCTPFLVALCTPAVYVTTIDENNILLDAQTAFFVSLAFNLLRPLNTLP 600
DB      541 KSAVLSAAGTFTWCTPFLVALCTPAVYVTTIDENNILLDAQTAFFVSLAFNLLRPLNTLP 600
QY      601 MVISSIVQASVSLKRLIFLSHEELEPDSIERRPVKDGGGTSITVRNATFTWASDPT 660
DB      601 MVISSIVQASVSLKRLIFLSHEELEPDSIERRPVKDGGGTSITVRNATFTWASDPT 660
QY      661 MVISSIVQASVSLKRLIFLSHEELEPDSIERRPVKDGGGTSITVRNATFTWASDPT 660
DB      661 MVISSIVQASVSLKRLIFLSHEELEPDSIERRPVKDGGGTSITVRNATFTWASDPT 660
QY      721 SLRENIILGCGQLEEFYRSVIOACALLPDLBITLPSGDRTEIGEKVNSGGCKORVSLAR 780
DB      721 SLRENIILGCGQLEEFYRSVIOACALLPDLBITLPSGDRTEIGEKVNSGGCKORVSLAR 780
QY      781 AVYSNADLYLPDDPLSAVDAAHVGKHI FENVIGPKKMLKNKTRILVTHSMSTLPQVDVILV 840
DB      781 AVYSNADLYLPDDPLSAVDAAHVGKHI FENVIGPKKMLKNKTRILVTHSMSTLPQVDVILV 840
QY      841 MSGKISMSGVYOELLARDGAFAEFLRTYASTEOBQDAEENGVTGVSQKGAKEKMGEM 900
DB      841 MSGKISMSGVYOELLARDGAFAEFLRTYASTEOBQDAEENGVTGVSQKGAKEKMGEM 900
QY      901 LVTBSAGQOLQRLSSSSSSYSQDISRHNSTAELQKAERKKEETKMLMEADQAQGOVYL 960
DB      901 LVTBSAGQOLQRLSSSSSSYSQDISRHNSTAELQKAERKKEETKMLMEADQAQGOVYL 960
```

```
QY      961 SVYWDYMKAIQGLFISFLSIFLPMCNHVSALASNYWLSLWTDPIVNGTOEHTKYRLSVYG 1020
DB      961 SVYWDYMKAIQGLFISFLSIFLPMCNHVSALASNYWLSLWTDPIVNGTOEHTKYRLSVYG 1020
QY      1021 ALGISOGIAVFGYMAVNSIGIILASRCLHYDLHSILRSMPSPREPTPSGULVNRFSKEL 1080
DB      1021 ALGISOGIAVFGYMAVNSIGIILASRCLHYDLHSILRSMPSPREPTPSGULVNRFSKEL 1080
QY      1081 DTVDSMIPEVYIKMMSGLFNVIGACIVILLATPIAIIIPPLGIYFFVOGRFYASSROL 1140
DB      1081 DTVDSMIPEVYIKMMSGLFNVIGACIVILLATPIAIIIPPLGIYFFVOGRFYASSROL 1140
QY      1141 KRLESVSRSPYSHENETLLGVSVIRAFEEQERFIHQSDLKVDENQAYYPSIVANRWLA 1200
DB      1141 KRLESVSRSPYSHENETLLGVSVIRAFEEQERFIHQSDLKVDENQAYYPSIVANRWLA 1200
QY      1201 VRLCEVNCIVLFPALFAVVISRHSLSAGVLGVSVYSLOVTTYLNMVLRMSSEMETNIVA 1260
DB      1201 VRLCEVNCIVLFPALFAVVISRHSLSAGVLGVSVYSLOVTTYLNMVLRMSSEMETNIVA 1260
QY      1261 VERLKEYSETEKAPMOIOETAPPSWPOVGRVFRNYCLYREDLPFLRHINVTINGG 1320
DB      1261 VERLKEYSETEKAPMOIOETAPPSWPOVGRVFRNYCLYREDLPFLRHINVTINGG 1320
QY      1321 EKVGIVRGTGAGKSLTGLFRINESABGEIIIDGINIAKIGLHDLRPKITIIPODPVLF 1380
DB      1321 EKVGIVRGTGAGKSLTGLFRINESABGEIIIDGINIAKIGLHDLRPKITIIPODPVLF 1380
QY      1381 SGIIRMLNDPPSOVSDEFWTSLFLAHLKDFVSLPKKLHECAGEGENISVGOEOLVCL 1440
DB      1381 SGIIRMLNDPPSOVSDEFWTSLFLAHLKDFVSLPKKLHECAGEGENISVGOEOLVCL 1440
QY      1441 ARALLRKTKILVDEEATAVVDLEETDDLIQSTIRFQFEDCVLTIAHLNTIMDYTRVIL 1500
DB      1441 ARALLRKTKILVDEEATAVVDLEETDDLIQSTIRFQFEDCVLTIAHLNTIMDYTRVIL 1500
QY      1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531
DB      1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531

RESULT 8
US-10-484-577-678
; Sequence 678, Application US/10484577
; Publication No. US20050032724A1
; GENERAL INFORMATION:
; APPLICANT: EPIDAUROS Biotechnologie Aktiengesellschaft
; TITLE OF INVENTION: Means and methods for improved treatment of cancer based on UGT1A
; FILE REFERENCE: P2285PCT-1
; CURRENT APPLICATION NUMBER: US/10/484,577
; CURRENT FILING DATE: 2004-01-22
; PRIOR APPLICATION NUMBER: PCT/EP 02/08220
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: EP 01 11 7608.8
; PRIOR FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: EP 02011710.7
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 683
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 678
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-484-577-678

Query Match      79.4%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MALRGFCGADSDPLMDMNNVTMTNSNPDFTKCFONTVLVWVPCFYLMACFPFFFLYLSRH 60
DB      1 MALRGFCGADSDPLMDMNNVTMTNSNPDFTKCFONTVLVWVPCFYLMACFPFFFLYLSRH 60
```

QY 61 DRGIOMTPLNKTALGFLMIMYCMADLFYSFWERSRGIFLAIPVLYSPPTLLGITTLLA 120
DB 61 DRGIOMTPLNKTALGFLMIMYCMADLFYSFWERSRGIFLAIPVLYSPPTLLGITTLLA 120
QY 121 TFLIQERRKGVSSGIMLTFMVALVICALILRSKIMTALKEBAQVDLFRDITFYVYFS 180
DB 121 TFLIQERRKGVSSGIMLTFMVALVICALILRSKIMTALKEBAQVDLFRDITFYVYFS 180
QY 181 LLLIQVLSCFSDSPFLSETIHDNPNCPRESSASFLSITITWMTTGLIVRGYROPLEGSD 240
DB 181 LLLIQVLSCFSDSPFLSETIHDNPNCPRESSASFLSITITWMTTGLIVRGYROPLEGSD 240
QY 241 LMSLNKEDTSEQVAVLYVKNMKKCAKTRKOPVAVVYSSKDPAPKSSSKYDAEVEAL 300
DB 241 LMSLNKEDTSEQVAVLYVKNMKKCAKTRKOPVAVVYSSKDPAPKSSSKYDAEVEAL 300
QY 301 IVKSPKEMWNSLFRKLYKTGPFYFLMSFPFKALHDLMPFSGPOLKLLIKFVNDTRAPD 360
DB 301 IVKSPKEMWNSLFRKLYKTGPFYFLMSFPFKALHDLMPFSGPOLKLLIKFVNDTRAPD 360
QY 361 WQGYFYVLLFVYKCLQTLVHQYFHICFVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420
DB 361 WQGYFYVLLFVYKCLQTLVHQYFHICFVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420
QY 421 GEIYNLMSVDAORFMDLATYINMTWSAPIOYIILALYILMTLIGSVYLAGVAVMTLMPYV 480
DB 421 GEIYNLMSVDAORFMDLATYINMTWSAPIOYIILALYILMTLIGSVYLAGVAVMTLMPYV 480
QY 481 AVNMAKTKTYOVAMKSKDNRIKLMNEILNGIKVLYAMELAFKDKVLAIRBELKYLK 540
DB 481 AVNMAKTKTYOVAMKSKDNRIKLMNEILNGIKVLYAMELAFKDKVLAIRBELKYLK 540
QY 541 KSAVLAVGFTFWCTPFLVALCTFAVYVITDENNIIDAQTAFAVSLAFNLRPPLNLP 600
DB 541 KSAVLAVGFTFWCTPFLVALCTFAVYVITDENNIIDAQTAFAVSLAFNLRPPLNLP 600
QY 601 MVISSIYOASVSLKRLAIFLSHELEPDSIERPVKCGGNGSTVNATFTWARSPT 660
DB 601 MVISSIYOASVSLKRLAIFLSHELEPDSIERPVKCGGNGSTVNATFTWARSPT 660
QY 661 LMGITFSPREGALVAVVGVCCKGSSLSALALAMDEKVEGVAIKGSVAVYPOQAMQND 720
DB 661 LMGITFSPREGALVAVVGVCCKGSSLSALALAMDEKVEGVAIKGSVAVYPOQAMQND 720
QY 721 SLRENIIFGQLEEPYRSYIQAICALIPDEILPESGRTEIGEGVNLSSGQKQVSLAR 780
DB 721 SLRENIIFGQLEEPYRSYIQAICALIPDEILPESGRTEIGEGVNLSSGQKQVSLAR 780
QY 781 AVYSNMDIYLPDPLSLAVDAHVGHIEENYIGPKGMKNTKRIILVTHSMGYLPQVDYIIV 840
DB 781 AVYSNMDIYLPDPLSLAVDAHVGHIEENYIGPKGMKNTKRIILVTHSMGYLPQVDYIIV 840
QY 841 MSGGKISEMGSYOELARDAFAEFLRTYASTEOBAEENGVTGVSQPKGAQOMENGM 900
DB 841 MSGGKISEMGSYOELARDAFAEFLRTYASTEOBAEENGVTGVSQPKGAQOMENGM 900
QY 901 LVTDSAGKQLOROQLSSSSSYSGDISRHHNSTAELQKAKEKFEWKLMLEADKAQTQGVKL 960
DB 901 LVTDSAGKQLOROQLSSSSSYSGDISRHHNSTAELQKAKEKFEWKLMLEADKAQTQGVKL 960
QY 961 SVYMDYKKAIGLFLSFSLIFLFCNHNYSALASNYWLSLWDDPVLNTOHTKTRLSVYG 1020
DB 961 SVYMDYKKAIGLFLSFSLIFLFCNHNYSALASNYWLSLWDDPVLNTOHTKTRLSVYG 1020
QY 1021 ALGISOGIAVFGYSMAVISGIIILASRCILAVDLHSLIRSPMSPEERTPSGNLVNRSFEL 1080
DB 1021 ALGISOGIAVFGYSMAVISGIIILASRCILAVDLHSLIRSPMSPEERTPSGNLVNRSFEL 1080
QY 1081 DTYDSMIPYIKNFMGSLENVIGACTIYLLATPIAIIIPPLGIYFFVORFYVASSROL 1140
DB 1081 DTYDSMIPYIKNFMGSLENVIGACTIYLLATPIAIIIPPLGIYFFVORFYVASSROL 1140
QY 1140 DTYDSMIPYIKNFMGSLENVIGACTIYLLATPIAIIIPPLGIYFFVORFYVASSROL 1140
DB 1140 DTYDSMIPYIKNFMGSLENVIGACTIYLLATPIAIIIPPLGIYFFVORFYVASSROL 1140

QY 1141 KRLSEVSRSPYSHFNELLGVSIVYRAFEOREFTHOSDLKVDENOKAYVPSIVANRMLA 1200
DB 1141 KRLSEVSRSPYSHFNELLGVSIVYRAFEOREFTHOSDLKVDENOKAYVPSIVANRMLA 1200
QY 1201 VRLCEVNCIVLPALFAVIRSRHSLAGVLGVSYSLOVTTYILNMLVRRMSEMETNIVA 1260
DB 1201 VRLCEVNCIVLPALFAVIRSRHSLAGVLGVSYSLOVTTYILNMLVRRMSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMOIOETAPPSWPQVGRVFNRYCLIRYREDDLPVLRIHINTNGG 1320
DB 1261 VERLKEYSETEKAPMOIOETAPPSWPQVGRVFNRYCLIRYREDDLPVLRIHINTNGG 1320
QY 1321 EKVIGVGTAGKSSSLTGLFRINESAGEITIIDGINIAKIGHDLAFKTIIPQDPVLF 1380
DB 1321 EKVIGVGTAGKSSSLTGLFRINESAGEITIIDGINIAKIGHDLAFKTIIPQDPVLF 1380
QY 1381 SGSLRNMULDPPSOYSDSEVWTSLELAHDKDFVSLPDKLHCEAGGENTSVGQROLVCL 1440
DB 1381 SGSLRNMULDPPSOYSDSEVWTSLELAHDKDFVSLPDKLHCEAGGENTSVGQROLVCL 1440
QY 1441 APALLRKTILVUDEATAVADLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRYVL 1500
DB 1441 APALLRKTILVUDEATAVADLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRYVL 1500
QY 1501 DKGEIOEYGA PSDILOQRGLFYSGAKDAGLV 1531
DB 1501 DKGEIOEYGA PSDILOQRGLFYSGAKDAGLV 1531

RESULT 9

US-10-889-503-19
Sequence 19, Application US/10889503
Publication No. US2005063968A1
GENERAL INFORMATION:
APPLICANT: Fox Chase Cancer Center
APPLICANT: Kruth, Gary D.
APPLICANT: Lee, Kun
APPLICANT: Belinsky, Martin G.
APPLICANT: Bain, Lisa J.
TITLE OF INVENTION: MRP-Related ABC Transporter Encoding
FILE REFERENCE: PCCC 98-02
CURRENT APPLICATION NUMBER: US/10/889,503
CURRENT FILING DATE: 2004-07-12
PRIOR APPLICATION NUMBER: US/09/647,140
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US99/06644
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 60/079,759
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/095,153
PRIOR FILING DATE: 1998-08-03
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 19
LENGTH: 1531
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-889-503-19

Query Match 79.4%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCASADGSDPLMDMNTWNTSNPDFTKCFQONTVLVWPCFYLMACPFYFLYSRH 60
DB 1 MALRGFCASADGSDPLMDMNTWNTSNPDFTKCFQONTVLVWPCFYLMACPFYFLYSRH 60
QY 61 DRGIOMTPLNKTALGFLMIMYCMADLFYSFWERSRGIFLAIPVLYSPPTLLGITTLLA 120
DB 61 DRGIOMTPLNKTALGFLMIMYCMADLFYSFWERSRGIFLAIPVLYSPPTLLGITTLLA 120
QY 121 TFLIQERRKGVSSGIMLTFMVALVICALILRSKIMTALKEBAQVDLFRDITFYVYFS 180
DB 121 TFLIQERRKGVSSGIMLTFMVALVICALILRSKIMTALKEBAQVDLFRDITFYVYFS 180

Db 121 TFLQLERRKGVSSGIMLTFMVALVCALALRSKIMTALKEADAQVDLFFDITFYVFS 180
Qy 181 LLLIQLVLSGSDSPLEFSETHDNPCESSASFLSRITFMWITGLIVRGYROPLEBSD 240
Db 181 LLLIQLVLSGSDSPLEFSETHDNPCESSASFLSRITFMWITGLIVRGYROPLEBSD 240
Qy 241 LMSLNKEDTSQVVPVIVKMKKECAKTRKQPVKVVVSSKDPAPKSSKVDANEVEAL 300
Db 241 LMSLNKEDTSQVVPVIVKMKKECAKTRKQPVKVVVSSKDPAPKSSKVDANEVEAL 300
Qy 301 IVKSPOKEMNSLTKULYKTRGPFYFLMSFFPKAIDHLMFSGPOLIKLLIFVNDYKAPD 360
Db 301 IVKSPOKEMNSLTKULYKTRGPFYFLMSFFPKAIDHLMFSGPOLIKLLIFVNDYKAPD 360
Qy 361 WOGIFYTVLTFVTAQCLQTLVHOYFHCIVSGMRKTAIVAGVYRKALVITNSARKSTV 420
Db 361 WOGIFYTVLTFVTAQCLQTLVHOYFHCIVSGMRKTAIVAGVYRKALVITNSARKSTV 420
Qy 421 GEIVNLMSVDAQRFMDLATYINMISAPLQVTLALYLLMLNIGPSVLAGVAVMYLMPVN 480
Db 421 GEIVNLMSVDAQRFMDLATYINMISAPLQVTLALYLLMLNIGPSVLAGVAVMYLMPVN 480
Qy 481 AVAMAKTKTVQVAMKSKDNRIKLMNELLNGIKYLKYAMELAPKDKYLAIRQEBELKYLK 540
Db 481 AVAMAKTKTVQVAMKSKDNRIKLMNELLNGIKYLKYAMELAPKDKYLAIRQEBELKYLK 540
Qy 541 KSAIYLSAVGTFWCTPFLVALCTFAVVTIDENNILDAQTAFAVSALFNILRPPLNLP 600
Db 541 KSAIYLSAVGTFWCTPFLVALCTFAVVTIDENNILDAQTAFAVSALFNILRPPLNLP 600
Qy 601 MVISSIVQASVSLKRLRIFLSHEELPEPSIERRPVKDGGTNSITVRNATFTWARSDBPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHEELPEPSIERRPVKDGGTNSITVRNATFTWARSDBPT 660
Qy 661 LMGITPSPBEALVAVVGQVCGKSSLLSALLAENDKRGSHVATKGSVAVYPOQAMIOND 720
Db 661 LMGITPSPBEALVAVVGQVCGKSSLLSALLAENDKRGSHVATKGSVAVYPOQAMIOND 720
Qy 721 SLRENIIFGCOLPEPYRSVIOACALPDLLELPISGDTIEGEGKVNLSGGOKORVSLAR 780
Db 721 SLRENIIFGCOLPEPYRSVIOACALPDLLELPISGDTIEGEGKVNLSGGOKORVSLAR 780
Qy 781 AVYSNADIYLPDDPLSAVDAHVKHIIFENVIGPKGMLNKTRILLVTHSMYLPQVDVIV 840
Db 781 AVYSNADIYLPDDPLSAVDAHVKHIIFENVIGPKGMLNKTRILLVTHSMYLPQVDVIV 840
Qy 841 MSGKISMSGYSQELIARDGAFAEFLRTYASTEQDAEBENGVTGVSQPGKEAKOMENGM 900
Db 841 MSGKISMSGYSQELIARDGAFAEFLRTYASTEQDAEBENGVTGVSQPGKEAKOMENGM 900
Qy 901 LVTHSAGKQOROLSSSSSSYSGDISRHNSSTAELOKAKKEBTMKLEADKAOTGOVKL 960
Db 901 LVTHSAGKQOROLSSSSSSYSGDISRHNSSTAELOKAKKEBTMKLEADKAOTGOVKL 960
Qy 961 SVYWDYMKAIQLFISFLSIFLMCMNVASALSNVWLSMTDPIVNGTOEHTKVRLSYVG 1020
Db 961 SVYWDYMKAIQLFISFLSIFLMCMNVASALSNVWLSMTDPIVNGTOEHTKVRLSYVG 1020
Qy 1021 ALGISQGIYAVFGYMAVSIIGIILASRCLHVDLHSILRSFMSFFERTPSGULVNFSEKL 1080
Db 1021 ALGISQGIYAVFGYMAVSIIGIILASRCLHVDLHSILRSFMSFFERTPSGULVNFSEKL 1080
Qy 1081 DTVOSMIPBEVIMKMFGLFENVIGACTIVLLATPRAAIIIPPLGILYFFVORPYVASSQOL 1140
Db 1081 DTVOSMIPBEVIMKMFGLFENVIGACTIVLLATPRAAIIIPPLGILYFFVORPYVASSQOL 1140
Qy 1141 KRLESVSRSPYVSHFNETLGVSVIRAREQOERFHQSDLKVDENQKAYYSIVANRMILA 1200
Db 1141 KRLESVSRSPYVSHFNETLGVSVIRAREQOERFHQSDLKVDENQKAYYSIVANRMILA 1200
Qy 1201 VRLFCVNCIYLPALPAVVISRHSISAGLVGLSVSYLQVTTYLNMVLVMSSEMETNIVA 1260
Db 1201 VRLFCVNCIYLPALPAVVISRHSISAGLVGLSVSYLQVTTYLNMVLVMSSEMETNIVA 1260

Db 1201 VRLFCVNCIYLPALPAVVISRHSISAGLVGLSVSYLQVTTYLNMVLVMSSEMETNIVA 1260
Qy 1261 VERLKESETEKEAPMOIOETAPSSWPQYGRVEFRNYCLRYREDLDFVLRHINVTINGG 1320
Db 1261 VERLKESETEKEAPMOIOETAPSSWPQYGRVEFRNYCLRYREDLDFVLRHINVTINGG 1320
Qy 1321 EKVGIVERTGAGSSLTGLIFRINESABGEIIGDINIAKIGLHDLRFKTIITIPQDPVLF 1380
Db 1321 EKVGIVERTGAGSSLTGLIFRINESABGEIIGDINIAKIGLHDLRFKTIITIPQDPVLF 1380
Qy 1381 SGLSRMNLDPFSQVSDSEWMTSLELAHLKDFVSALEPKLHECAGGENLSVGOROLVCL 1440
Db 1381 SGLSRMNLDPFSQVSDSEWMTSLELAHLKDFVSALEPKLHECAGGENLSVGOROLVCL 1440
Qy 1441 ABALIRKTKILVDEATAVADLETDLLIOSTIRTOPEDCTVLTIANRIANTIMDYTRYVL 1500
Db 1441 ABALIRKTKILVDEATAVADLETDLLIOSTIRTOPEDCTVLTIANRIANTIMDYTRYVL 1500
Qy 1501 DKGEIOEYGAUSDLLQORGLFYSMAKDAGLV 1531
Db 1501 DKGEIOEYGAUSDLLQORGLFYSMAKDAGLV 1531

RESULT 10
US-10-756-149-5033
; Sequence 5033, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Azit, Natasha
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5033
; LENGTH: 1531
; TYPE: PRM
; ORGANISM: Homo Sapiens
US-10-756-149-5033

Query Match 79.4%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRGFCSADGSDPLMDMNTVMTNSNDPTKCFQNTVLLVWPCTYLMACFPFFLYLSRH 60
Db 1 MALRGFCSADGSDPLMDMNTVMTNSNDPTKCFQNTVLLVWPCTYLMACFPFFLYLSRH 60
Qy 61 DRGYIOMTPLNKTATLGFLLMIVCMADLFYSFMRSGRFLAVPLVSPPLIGITTLA 120
Db 61 DRGYIOMTPLNKTATLGFLLMIVCMADLFYSFMRSGRFLAVPLVSPPLIGITTLA 120
Qy 121 TFLQLERRKGVSSGIMLTFMVALVCALALRSKIMTALKEADAQVDLFFDITFYVFS 180
Db 121 TFLQLERRKGVSSGIMLTFMVALVCALALRSKIMTALKEADAQVDLFFDITFYVFS 180
Qy 181 LLLIQLVLSGSDSPLEFSETHDNPCESSASFLSRITFMWITGLIVRGYROPLEBSD 240
Db 181 LLLIQLVLSGSDSPLEFSETHDNPCESSASFLSRITFMWITGLIVRGYROPLEBSD 240
Qy 241 LMSLNKEDTSQVVPVIVKMKKECAKTRKQPVKVVVSSKDPAPKSSKVDANEVEAL 300
Db 241 LMSLNKEDTSQVVPVIVKMKKECAKTRKQPVKVVVSSKDPAPKSSKVDANEVEAL 300
Qy 301 IVKSPOKEMNSLTKULYKTRGPFYFLMSFFPKAIDHLMFSGPOLIKLLIFVNDYKAPD 360
Db 301 IVKSPOKEMNSLTKULYKTRGPFYFLMSFFPKAIDHLMFSGPOLIKLLIFVNDYKAPD 360
Qy 361 WOGIFYTVLTFVTAQCLQTLVHOYFHCIVSGMRKTAIVAGVYRKALVITNSARKSTV 420
Db 361 WOGIFYTVLTFVTAQCLQTLVHOYFHCIVSGMRKTAIVAGVYRKALVITNSARKSTV 420

Db 361 WQGYFTVLLFVTACLOTVLHQYFHLCEVSGMRKTAIVIGAVYRKALVITNSARKSTV 420
Qy 421 GEIYNLMSVPAQRPMDLATYINMTMSAPLOVITLALYLLMLNGSVLAGVAVMLAMPVN 480
Db 421 GEIYNLMSVPAQRPMDLATYINMTMSAPLOVITLALYLLMLNGSVLAGVAVMLAMPVN 480
Qy 481 AVMAKTKTYOVAMHMSKDNRIKLMNEILNGIKYLKLYAMELAKDKVLAIRQBELKYLK 540
Db 481 AVMAKTKTYOVAMHMSKDNRIKLMNEILNGIKYLKLYAMELAKDKVLAIRQBELKYLK 540
Qy 541 KSAVLASVGTFTWCTPFLVALCTFAVYVITDENNNILDAQAFVSLAFNLIRFPLNLP 600
Db 541 KSAVLASVGTFTWCTPFLVALCTFAVYVITDENNNILDAQAFVSLAFNLIRFPLNLP 600
Qy 601 MVLSIIYQASVSLKRLNIFLSHELEPDSIRRRPVYKGGGNGSTVYNNATPTAKRSPPPT 660
Db 601 MVLSIIYQASVSLKRLNIFLSHELEPDSIRRRPVYKGGGNGSTVYNNATPTAKRSPPPT 660
Qy 661 LNCITFSPREGALVAVVGVCGCKSSLSLALLAEMDKVGHVATKGSVAVYPOQAMIOND 720
Db 661 LNCITFSPREGALVAVVGVCGCKSSLSLALLAEMDKVGHVATKGSVAVYPOQAMIOND 720
Qy 721 SLRENILFGCQLEBPYRSVIOACALLPDLEILPSGDRTEIGEKVNLSGGQKQVSLAR 780
Db 721 SLRENILFGCQLEBPYRSVIOACALLPDLEILPSGDRTEIGEKVNLSGGQKQVSLAR 780
Qy 781 AVYSNADIIYFDDLSAVDAVGHVIGHIVNVIKPGKMLKNKTRILVTHSMSTLPQVDYIIV 840
Db 781 AVYSNADIIYFDDLSAVDAVGHVIGHIVNVIKPGKMLKNKTRILVTHSMSTLPQVDYIIV 840
Qy 841 MSGGKISEMGSYOELLARDGFAFELRTYASTEOEDABENGVTGVSQPGKEAKOMENGM 900
Db 841 MSGGKISEMGSYOELLARDGFAFELRTYASTEOEDABENGVTGVSQPGKEAKOMENGM 900
Qy 901 LVYDSAGKOLQROLSSSSSYSGDISRHHNSTAEIOKAKEKEETWKLMEADKATQGVKL 960
Db 901 LVYDSAGKOLQROLSSSSSYSGDISRHHNSTAEIOKAKEKEETWKLMEADKATQGVKL 960
Qy 961 SVYWDYKAKGLFISFLIFPMCNHVSALASNYWLSMTDDPIVNGTOEHTKRLSVYG 1020
Db 961 SVYWDYKAKGLFISFLIFPMCNHVSALASNYWLSMTDDPIVNGTOEHTKRLSVYG 1020
Qy 1021 ALGISOGIAVFGVMAVSIIGIILASRCILHVDLHSIIRSPMSFEERPPSGNLVRFSEKL 1080
Db 1021 ALGISOGIAVFGVMAVSIIGIILASRCILHVDLHSIIRSPMSFEERPPSGNLVRFSEKL 1080
Qy 1081 DTVDSMIPVYIKPMGSLFNVIGACIYILLATPIAIIIPPLGIYFVQRFVYASSROL 1140
Db 1081 DTVDSMIPVYIKPMGSLFNVIGACIYILLATPIAIIIPPLGIYFVQRFVYASSROL 1140
Qy 1141 KRLSVRSRPVYSHFNFTLLIGSVYTRAPEROERFIHQSDLKVDENQKAYPSIVANWMLA 1200
Db 1141 KRLSVRSRPVYSHFNFTLLIGSVYTRAPEROERFIHQSDLKVDENQKAYPSIVANWMLA 1200
Qy 1201 VRLCEVNCIYLPALFAVISRHSLSAGLVGLSVSYLQVTTYLNMVLRMSSEMETNIVA 1260
Db 1201 VRLCEVNCIYLPALFAVISRHSLSAGLVGLSVSYLQVTTYLNMVLRMSSEMETNIVA 1260
Qy 1261 VERLKEVSETEKAPMOIOETAPPSWPQVGRVFRNYCJRYREDDLPVLRHINVTNGG 1320
Db 1261 VERLKEVSETEKAPMOIOETAPPSWPQVGRVFRNYCJRYREDDLPVLRHINVTNGG 1320
Qy 1321 EKVGIVERTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
Db 1321 EKVGIVERTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
Qy 1381 SSGSRNMLDPPSOYSDSEWVTSLEIATKDFVSLPKLHECAGEGEMNSVQGRQVYCL 1440
Db 1381 SSGSRNMLDPPSOYSDSEWVTSLEIATKDFVSLPKLHECAGEGEMNSVQGRQVYCL 1440
Qy 1441 ARALLRKTKLIVLDEATAVADLETDLLIOSTIRTOFEDCTVLTAAHLANTIMDYTRYIVL 1500
Db 1441 ARALLRKTKLIVLDEATAVADLETDLLIOSTIRTOFEDCTVLTAAHLANTIMDYTRYIVL 1500

Qy 1501 DKGEIOEYGAPSDLLQORGLFYSGAKDAGLV 1531
Db 1501 DKGEIOEYGAPSDLLQORGLFYSGAKDAGLV 1531

RESULT 11
US-09-939-853A-87
; Sequence 87, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939, 853A
; PRIOR APPLICATION NUMBER: 2001-08-27
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 1515
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-87

Query Match 78.5%; Score 7769; DB 3; Length 1515;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1515; Conservative 0; Mismatches 0; Indels 0; Caps 0;

Qy 17 DNVVNTNSPDDTKPCQNVLVVWVPCFYLMACPFYLYLSRHDGCIQMTPLNKTKTA 76
Db 1 DNVVNTNSPDDTKPCQNVLVVWVPCFYLMACPFYLYLSRHDGCIQMTPLNKTKTA 60

Qy 77 LGFLMLVCAADLFYSEWERSRGIPLAPVPLVSGPTLLGTTLLATFLIQLERRKGVQSSG 136
Db 61 LGFLMLVCAADLFYSEWERSRGIPLAPVPLVSGPTLLGTTLLATFLIQLERRKGVQSSG 120

Qy 137 IMTFMVLVLCALILRSKIMTALKEDAQVDLFRDITFYVYFSLLIQVLISCFSDRSP 196
Db 121 IMTFMVLVLCALILRSKIMTALKEDAQVDLFRDITFYVYFSLLIQVLISCFSDRSP 180

Qy 197 LFSETHDHPCESSASFLSRTTFPMWITGLYRGYRQPLEGSDLMSLNKEDTSEOVVPV 256
Db 181 LFSETHDHPCESSASFLSRTTFPMWITGLYRGYRQPLEGSDLMSLNKEDTSEOVVPV 240

Qy 257 LVNWMKKECAKTKQPKVYVYSSKDPAPQKSSKVDANBEVLAIVKSPQKEMNPISLFKY 316
Db 241 LVNWMKKECAKTKQPKVYVYSSKDPAPQKSSKVDANBEVLAIVKSPQKEMNPISLFKY 300

Qy 317 LYXTPGYPFLMSPFPKAIHDMFSGPOILIKLIFKVNDRKADWQGYFYTVLLPYTACL 376
Db 301 LYXTPGYPFLMSPFPKAIHDMFSGPOILIKLIFKVNDRKADWQGYFYTVLLPYTACL 360

Qy 377 QTLVHLQYPRICVSGMRITKTVAVYRKALVITNSARKSSVYGEIVNIMSDAQRFMD 436
Db 361 QTLVHLQYPRICVSGMRITKTVAVYRKALVITNSARKSSVYGEIVNIMSDAQRFMD 420

Qy 437 LATYINMISAPQVILALYLLMLNGPSVLAGVAVVLMVFNVAWMAKTKTYQVAAHMK 496
Db 421 LATYINMISAPQVILALYLLMLNGPSVLAGVAVVLMVFNVAWMAKTKTYQVAAHMK 480

Qy 497 SKONRIKLMNEILNGIKYLKLYAMELAFKDKVLAIRQBELKYLKKSAYIASAVGTFWVCT 556
Db 481 SKONRIKLMNEILNGIKYLKLYAMELAFKDKVLAIRQBELKYLKKSAYIASAVGTFWVCT 540

Qy 557 PFLVALCTFAVYVITDENNNILDAQFVSLAFNLIRFPLNLPVYISSIVQASVSLKRL 616

|||||
Db 541 PFLVALCTFAVYVITIDENNIIIDAQTAFAVSALFNILRPLNILPMVSISSIVQASVSLKRL 600
Qy 617 RIFLSHEBELPDSIERRPVKDGGGNSITVNNATFTWARSPTLNGITFSIPGALVAV 676
Db 601 RIFLSHEBELPDSIERRPVKDGGGNSITVNNATFTWARSPTLNGITFSIPGALVAV 660
Qy 677 VGVGCGKSSLSLSALLAEMDKVEGHVAKGSVAVYPOQAWIQNDSLRENILFGCQLEBPY 736
Db 661 VGVGCGKSSLSLSALLAEMDKVEGHVAKGSVAVYPOQAWIQNDSLRENILFGCQLEBPY 720
Qy 737 YRSYIQAACALLPDLIELPSGDRTEIGEKNVLSGGQKORVSLARAAYSNADIYLPDPLS 796
Db 721 YRSYIQAACALLPDLIELPSGDRTEIGEKNVLSGGQKORVSLARAAYSNADIYLPDPLS 780
Qy 797 AYDAHVGHKHFENYIGPKGMKNKTRILVTSHMSLTPVDVIVMSGKISEMSYQEL 856
Db 781 AYDAHVGHKHFENYIGPKGMKNKTRILVTSHMSLTPVDVIVMSGKISEMSYQEL 840
Qy 857 ARDGAFAEFLRTVASTEOEODAEENGVTGVSQPKAKOMENGMVLTDSAGKOLQROUSS 916
Db 841 ARDGAFAEFLRTVASTEOEODAEENGVTGVSQPKAKOMENGMVLTDSAGKOLQROUSS 900
Qy 917 SSSYSGDISRHHNSTAEIQAKAKEETWKIMADKAOTGOVKLSVYWDYMKALGLFISF 976
Db 901 SSSYSGDISRHHNSTAEIQAKAKEETWKIMADKAOTGOVKLSVYWDYMKALGLFISF 960
Qy 977 LSTFLPMONHVSALASNYWLSLMTDDPIVNGTOHETKRLSVYGLGISQGIIVAGVSYMA 1036
Db 961 LSTFLPMONHVSALASNYWLSLMTDDPIVNGTOHETKRLSVYGLGISQGIIVAGVSYMA 1020
Qy 1037 VSIGIILASRCLAHYDLHLSIRSPMSPEPTPSGNLVNRFSEKEDTVDMSIPEVIMKFMG 1096
Db 1021 VSIGIILASRCLAHYDLHLSIRSPMSPEPTPSGNLVNRFSEKEDTVDMSIPEVIMKFMG 1080
Qy 1097 SLFNVIGACIYIILATPIAIIIPPLGLIYFPVORFYVASSRQKRLSVSRSPYSHFN 1156
Db 1081 SLFNVIGACIYIILATPIAIIIPPLGLIYFPVORFYVASSRQKRLSVSRSPYSHFN 1140
Qy 1157 ETLIGVSVIRAFERBERFIHOSDLKVDENOKAYPSIVANRWLAVRLCVCNCTVLPAL 1216
Db 1141 ETLIGVSVIRAFERBERFIHOSDLKVDENOKAYPSIVANRWLAVRLCVCNCTVLPAL 1200
Qy 1217 FAVISRHSLSAGLVGLSVYSLSQVTTYLNMLVRMSSEMETNI VAVERLKEVSETEKEAPW 1276
Db 1201 FAVISRHSLSAGLVGLSVYSLSQVTTYLNMLVRMSSEMETNI VAVERLKEVSETEKEAPW 1260
Qy 1277 QIOETAPSSWPQVGRVFEFRNYCLARYREDLDFVLHINVTINGEKVGI VGTGAGKSSL 1336
Db 1261 QIOETAPSSWPQVGRVFEFRNYCLARYREDLDFVLHINVTINGEKVGI VGTGAGKSSL 1320
Qy 1337 TLGFRINESAGEIIIDIGINIAKIGLHDLRKTITIIODPVLFGSILRMLNDPESQYSD 1396
Db 1321 TLGFRINESAGEIIIDIGINIAKIGLHDLRKTITIIODPVLFGSILRMLNDPESQYSD 1380
Qy 1397 BEVWTSLELAHKDFVSALPKLDHECABGGENLSVGOQOLVCLARALLRKTIIIVLDEA 1456
Db 1381 BEVWTSLELAHKDFVSALPKLDHECABGGENLSVGOQOLVCLARALLRKTIIIVLDEA 1440
Qy 1457 TAAVDELFTDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAESDILQ 1516
Db 1441 TAAVDELFTDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAESDILQ 1500
Qy 1517 ORGLFYSMAXDAGLV 1531
Db 1501 ORGLFYSMAXDAGLV 1515

RESULT 12

US-10-618-281-42

; Sequence 42, Application US/10618281

; Publication No. US20040219609A1

; GENERAL INFORMATION:

APPLICANT: Day, Anthony G.
APPLICANT: Estell, David A.
APPLICANT: Lyons, Eric H.
APPLICANT: Yao, Jian
TITLE OR INVENTION: Methods for Modulating Proteins Not
TITLE OR INVENTION: Previously known as Proteases
FILE REFERENCE: GC773-2
CURRENT APPLICATION NUMBER: US/10/618, 281
CURRENT FILING DATE: 2003-07-11
PRIOR APPLICATION NUMBER: US 60/395,325
PRIOR FILING DATE: 2002-07-12
NUMBER OF SEQ ID NOS: 92
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 42
LENGTH: 1515
TYPE: PR
ORGANISM: Homo sapiens
US-10-618-281-42

Query Match 78.4%; Score 7763; DB 5; Length 1515;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1514; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 17 DNANTNTSNPDFKTCQNTYLVWVPCFYLMACPPFYLISSRDGTYQMTPLNKTKTA 76
Db 1 DNANTNTSNPDFKTCQNTYLVWVPCFYLMACPPFYLISSRDGTYQMTPLNKTKTA 60
Qy 77 LGFLMIVCANMDELYSFWERSRGIFLAPVFLVSPTLGITLTLATFLIQLERRKGVOSG 136
Db 61 LGFLMIVCANMDELYSFWERSRGIFLAPVFLVSPTLGITLTLATFLIQLERRKGVOSG 120
Qy 137 IMLFMWLVALCALAIRSKIMTALKSDAQVDLFRDITFYVYFSLLIQVLSCFSDRSP 196
Db 121 IMLFMWLVALCALAIRSKIMTALKSDAQVDLFRDITFYVYFSLLIQVLSCFSDRSP 180
Qy 197 LFSETHIDPNCPRESSASFLSRITFWWITGLIYVGRQOPLSGSLMSLNKEDTSEQVVPV 256
Db 181 LFSETHIDPNCPRESSASFLSRITFWWITGLIYVGRQOPLSGSLMSLNKEDTSEQVVPV 240
Qy 257 LYKMKKCEKATROPKOPKVVYSSKDPAPKSSKYDANEVEALIVKSPKEMNPSLPKV 316
Db 241 LYKMKKCEKATROPKOPKVVYSSKDPAPKSSKYDANEVEALIVKSPKEMNPSLPKV 300
Qy 317 LYKTFGPFLMSPEPKAIHDLMFSGPOLKLIKFVNDTKAPDMOGYFTVLLFVTAQL 376
Db 301 LYKTFGPFLMSPEPKAIHDLMFSGPOLKLIKFVNDTKAPDMOGYFTVLLFVTAQL 360
Qy 377 QTLVLHOFYFHCYFSGNRIKTAIVGAYYRKALVITNSARKSSTVGEIYNLMSVDAQRMD 436
Db 361 QTLVLHOFYFHCYFSGNRIKTAIVGAYYRKALVITNSARKSSTVGEIYNLMSVDAQRMD 420
Qy 437 LATYINMWSAPLOVITLALYLMNLGSPYLAGVAVWVWLVNANVAMKTKTYOVAMK 496
Db 421 LATYINMWSAPLOVITLALYLMNLGSPYLAGVAVWVWLVNANVAMKTKTYOVAMK 480
Qy 497 SKDRIKLMNLIINGIKVLKLYAMELAFKDKVLAIRBELVYLLKKSAYLSAVGTFTWCT 556
Db 481 SKDRIKLMNLIINGIKVLKLYAMELAFKDKVLAIRBELVYLLKKSAYLSAVGTFTWCT 540
Qy 557 PFLVALCTFAVYVITIDENNIIIDAQTAFAVSALFNILRPLNILPMVSISSIVQASVSLKRL 616
Db 541 PFLVALCTFAVYVITIDENNIIIDAQTAFAVSALFNILRPLNILPMVSISSIVQASVSLKRL 600
Qy 617 RIFLSHEBELPDSIERRPVKDGGGNSITVNNATFTWARSPTLNGITFSIPGALVAV 676
Db 601 RIFLSHEBELPDSIERRPVKDGGGNSITVNNATFTWARSPTLNGITFSIPGALVAV 660
Qy 677 VGVGCGKSSLSLSALLAEMDKVEGHVAKGSVAVYPOQAWIQNDSLRENILFGCQLEBPY 736
Db 661 VGVGCGKSSLSLSALLAEMDKVEGHVAKGSVAVYPOQAWIQNDSLRENILFGCQLEBPY 720
Qy 737 YRSYIQAACALLPDLIELPSGDRTEIGEKNVLSGGQKORVSLARAAYSNADIYLPDPLS 796

Db 721 YRSVIAQACALLPDLLEILPSGDRTEIGEKVNLSCGQKQKRVSLARAVYSNADIVLFDPLS 780
Qy 797 AVDAHVGHKIPENYIGPKMKKNTRLIVTHSMGYLPQVDVYIIMSGGKISEMGSYOELL 856
Db 781 AVDAHVGHKIPENYIGPKMKKNTRLIVTHSMGYLPQVDVYIIMSGGKISEMGSYOELL 840
Qy 857 ARDAFAEFLRTVASTEOBDAENGVTGVSQPKKAKOMENGMVLTDSAGKOLQROUSS 916
Db 841 ARDAFAEFLRTVASTEOBDAENGVTGVSQPKKAKOMENGMVLTDSAGKOLQROUSS 900
Qy 917 SSSYSGQISRHNSHTABLOKAEAKKETWKLMEADKQOTGOVKLSVTYWDYKAIQGLPISF 976
Db 901 SSSYSGQISRHNSHTABLOKAEAKKETWKLMEADKQOTGOVKLSVTYWDYKAIQGLPISF 960
Qy 977 LSTFLPMONVSAASNYWLSLMTDDPIVNGTOEHTKVRLSVYALGISQIAVFGSMA 1036
Db 961 LSTFLPMONVSAASNYWLSLMTDDPIVNGTOEHTKVRLSVYALGISQIAVFGSMA 1020
Qy 1037 VSIGGILASRCLHVDLHLSILRSPMSFERTPSGNLVNRFSEKLDVTDSMTPEVIKMPMG 1096
Db 1021 VSIGGILASRCLHVDLHLSILRSPMSFERTPSGNLVNRFSEKLDVTDSMTPEVIKMPMG 1080
Qy 1097 SLENVIGACIVILATPIAIIIPPLGLIYFVQRFVYASSRQKRLSEVSRSFVYSHFN 1156
Db 1081 SLENVIGACIVILATPIAIIIPPLGLIYFVQRFVYASSRQKRLSEVSRSFVYSHFN 1140
Qy 1157 ETLIGVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWMLAVRLECVGNCIVLPAAL 1216
Db 1141 ETLIGVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRWMLAVRLECVGNCIVLPAAL 1200
Qy 1217 FAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSSEMETNIIVABRLKEXSETEKAPW 1276
Db 1201 FAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSSEMETNIIVABRLKEXSETEKAPW 1260
Qy 1277 QIOETAPSSMPQYGRVGFNRYCLRYREDLDFVLRHINVTINGEKVIGVGTGAGKSL 1336
Db 1261 QIOETAPSSMPQYGRVGFNRYCLRYREDLDFVLRHINVTINGEKVIGVGTGAGKSL 1320
Qy 1337 TLGIFRINESAEGIIIDGINIAKIGLHDLRFKTTIIPQDPVLFSSGLRNMLDPPSOYSD 1396
Db 1321 TLGIFRINESAEGIIIDGINIAKIGLHDLRFKTTIIPQDPVLFSSGLRNMLDPPSOYSD 1380
Qy 1397 BEWVTSLELAHLKDFVAFALPDKLDHECAEGENLSVQOROLVCLARALLRKTILVDEA 1456
Db 1381 BEWVTSLELAHLKDFVAFALPDKLDHECAEGENLSVQOROLVCLARALLRKTILVDEA 1440
Qy 1457 TAAVDETDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQ 1516
Db 1441 TAAVDETDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQ 1500
Qy 1517 QRGIFYSMADAGLV 1531
Db 1501 QRGIFYSMADAGLV 1515

RESULT 13

US-10-408-765A-1718
; Sequence 1718, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Bojin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Marnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408.765A
; NUMBER OF SEQ ID NOS: 3077

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1718
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1718
Query Match 71.1%; Score 7037; DB 4; Length 1388;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1387; Conservative 0; Mismatches 1; Indels 42; Gaps 1;
Qy 67 MTPLNKTKTALGFLMIWCVNADLIFYSPWERSRGIFLAPVFLVSEFTLITLTLATFLIQL 126
Db 1 MTPLNKTKTALGFLMIWCVNADLIFYSPWERSRGIFLAPVFLVSEFTLITLTLATFLIQL 60
Qy 127 ERRKVOSSGIMLTFMVLVLCALAIIRSKIMTALKEADOVLFRDITFYVYFSLLIQL 186
Db 61 ERRKVOSSGIMLTFMVLVLCALAIIRSKIMTALKEADOVLFRDITFYVYFSLLIQL 97
Qy 187 VLSGFSRSLPFSRTIHDPNCPRESSASFSLRTTFMWTGLIVRGYQPLEGSDMLNKL 246
Db 98 -----NPCSSASFSLRTTFMWTGLIVRGYQPLEGSDMLNKL 138
Qy 247 EDTSEGVVPLVKNMKKECAKTRKQPYKVYSSKDPAPQKSSKVDANEVEALIVSPQ 306
Db 139 EDTSEGVVPLVKNMKKECAKTRKQPYKVYSSKDPAPQKSSKVDANEVEALIVSPQ 198
Qy 307 KEMNPSEFKVLYTFCGYFLMSFFPKAIDHLMFSGQIIKLKIFVNDTKAPDMQGYFY 366
Db 199 KEMNPSEFKVLYTFCGYFLMSFFPKAIDHLMFSGQIIKLKIFVNDTKAPDMQGYFY 258
Qy 367 TULLFTYACLOTVLHOYFHIQVSGMRITAVYGAAYRRALVTNSARSSVTGETVNL 426
Db 259 TULLFTYACLOTVLHOYFHIQVSGMRITAVYGAAYRRALVTNSARSSVTGETVNL 318
Qy 427 MSYDAQRFMDLATYINMIWAPLOVILATYLMNLGSPSYLAGVAVWLVVPAVNAWAMK 486
Db 319 MSYDAQRFMDLATYINMIWAPLOVILATYLMNLGSPSYLAGVAVWLVVPAVNAWAMK 378
Qy 487 TKTYOVAHMSKNRILKAMEIINGIKVLKLYAMELAFKDKVLAIRQEBLVKKSAVLS 546
Db 379 TKTYOVAHMSKNRILKAMEIINGIKVLKLYAMELAFKDKVLAIRQEBLVKKSAVLS 438
Qy 547 ANGTFTWVCTPPLVALCTFAVYVTTIDENNLIDQOTAFVSLATNIIIRFPLNIPMTYSSI 606
Db 439 ANGTFTWVCTPPLVALCTFAVYVTTIDENNLIDQOTAFVSLATNIIIRFPLNIPMTYSSI 498
Qy 607 VQASVSLKRLRIFLSHELEPDSIERRPVQGGGINSITVRNATFTWARSDDPTLNGITF 666
Db 499 VQASVSLKRLRIFLSHELEPDSIERRPVQGGGINSITVRNATFTWARSDDPTLNGITF 558
Qy 667 SIEGALVAVVGVGCGKSSLSALLAEMDKVGHVAKGSVAVVPQAMVIONDSIRENI 726
Db 559 SIEGALVAVVGVGCGKSSLSALLAEMDKVGHVAKGSVAVVPQAMVIONDSIRENI 618
Qy 727 LFGCQIBEPYRVSIVQACALLPDLLEILPSGDRTEIGEKVNLSCGQKQKRVSLARAVYSNA 786
Db 619 LFGCQIBEPYRVSIVQACALLPDLLEILPSGDRTEIGEKVNLSCGQKQKRVSLARAVYSNA 678
Qy 787 DITLFDPPISAVAHVGKHI FENVIGPKMKKNTRLIVTHSMGYLPQVDVYIIMSGGKI 846
Db 679 DITLFDPPISAVAHVGKHI FENVIGPKMKKNTRLIVTHSMGYLPQVDVYIIMSGGKI 738
Qy 847 SEMGSYQELLARDGAFAEFLRTVASTEOBDAENGVTGVSQPKKAKOMENGMVLTDSA 906
Db 739 SEMGSYQELLARDGAFAEFLRTVASTEOBDAENGVTGVSQPKKAKOMENGMVLTDSA 798
Qy 907 GKOLQOLSSSSSYSDISRHNSHTABLOKAEAKKETWKLMEADKQOTGOVKLSVTYWDY 966
Db 799 GKOLQOLSSSSSYSDISRHNSHTABLOKAEAKKETWKLMEADKQOTGOVKLSVTYWDY 858
Qy 967 MKAIGLFIPLSLIFLMONVSAASNYWLSLMTDDPIVNGTOEHTKVRLSVYALGISQ 1026
Db 967 MKAIGLFIPLSLIFLMONVSAASNYWLSLMTDDPIVNGTOEHTKVRLSVYALGISQ 1026

Db 859 MKAIGLPISEIIFLFCMCHVNSALASNYMLWTDDPIVNGTQEHTKVRLSVYGAIGISQ 918
Qy 1027 GIAVFGSMAYSIGIILASRCLHVDLHLSILRSPMSFFERTPSGNLVNRFSEKELDTVDSM 1086
Db 919 GIAVFGSMAYSIGIILASRCLHVDLHLSILRSPMSFFERTPSGNLVNRFSEKELDTVDSM 978
Qy 1087 IPEYIKMFMGSLFENVIGACIYIILATPIAAIIPPLGLIYFFVQRFYVASSRQIKRLSEV 1146
Db 979 IPEYIKMFMGSLFENVIGACIYIILATPIAAIIPPLGLIYFFVQRFYVASSRQIKRLSEV 1038
Qy 1147 SRSVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIYANMVLAVRLECV 1206
Db 1039 SRSVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIYANMVLAVRLECV 1098
Qy 1207 GNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYANMLVRMSSEMETNIIVAVERLKE 1266
Db 1099 GNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYANMLVRMSSEMETNIIVAVERLKE 1158
Qy 1267 YSETEKEAPMOIQETAPPSSWPQVGRVFRNVCIRYREDLDPVLRIHVVTTINGEKEGIV 1326
Db 1159 YSETEKEAPMOIQETAPPSSWPQVGRVFRNVCIRYREDLDPVLRIHVVTTINGEKEGIV 1218
Qy 1327 GRTAGKSSLTGLFRINESAGEIIDGINIAKIGLHDLFKITIIPODPVLRSGSLRM 1386
Db 1219 GRTAGKSSLTGLFRINESAGEIIDGINIAKIGLHDLFKITIIPODPVLRSGSLRM 1278
Qy 1387 NLDPFSQYSDDEWVTSLELAHLKDFVSALPKDLDEHCAEGENLSVGOQROVLCLARALLR 1446
Db 1279 NLDPFSQYSDDEWVTSLELAHLKDFVSALPKDLDEHCAEGENLSVGOQROVLCLARALLR 1338
Qy 1447 KTKILVLDEATPAVDLETFDDLIQSTIRQPEDCTVLTIAHRLNTIMDYTR 1496
Db 1339 KTKILVLDEATPAVDLETFDDLIQSTIRQPEDCTVLTIAHRLNTIMDYTR 1388

RESULT 14
US-09-939-853A-88
; Sequence 88, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939, 853A
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/268,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 88
; LENGTH: 1528
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-939-853A-88

Query Match 70.7%; Score 7002.5; DB 3; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

Qy 1 MALRFGCAGDSDDLMDNNVTWNTISNPDFTKCFQNTVIVWPCFYLAACPPFYLYLSRH 60
Db 1 MALRFGCAGDSDDLMDNNVTWNTISNPDFTKCFQNTVIVWPCFYLAACPPFYLYLSRH 60
Qy 61 DRGYQWMLNKTKALGFLIMIVCMADLFVSFMRSRGIFLAPFLVSPITLGIITLLA 120
Db 61 DRGYQWMLNKTKALGFLIMIVCMADLFVSFMRSRGIFLAPFLVSPITLGIITLLA 120

Qy 121 TFLIQERRRGVSSGIMLTFMWLVALYCALAALRSKMTALKEDAOVDLFEDITFYFYS 180
Db 121 TFLIQERRRGVSSGIMLTFMWLVALYCALAALRSKMTALKEDAOVDLFEDITFYFYS 180
Qy 181 LLLIQVLSGCSGDSPLFSETIHDNPNCPSSSASFSLRIFPMWITGLIVGGRPLGSD 240
Db 181 LLLIQVLSGCSGDSPLFSETIHDNPNCPSSSASFSLRIFPMWITGLIVGGRPLGSD 240
Qy 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKNVYS-SKDPAPQRESSKVDANEVEYA 299
Db 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKNVYS-SKDPAPQRESSKVDANEVEYA 300
Qy 300 LTVSPOKEMNPSEKVLKTTTGGPFLMSFFPKAIDHLMPSGQIILKLIKPNDTKAP 359
Db 301 LTVSPOKEMNPSEKVLKTTTGGPFLMSFFPKAIDHLMPSGQIILKLIKPNDTKAP 360
Qy 360 DMQGYFTYALFVSACIQTLALHQYFHI CFVSGKRITAVAGAYRRALLITNARSST 420
Db 361 DMQGYFTYALFVSACIQTLALHQYFHI CFVSGKRITAVAGAYRRALLITNARSST 420
Qy 420 VGEIVNLMSYDAQRFMDLATYINMWSAPLOVILATYLMNLGPSYLAGVAVVWLVVPV 479
Db 421 VGEIVNLMSYDAQRFMDLATYINMWSAPLOVILATYLMNLGPSYLAGVAVVWLVVPV 480
Qy 480 NAYMAKTKTYQVAHMSKONRIKLANEILNGIKVLYKYAMELAFKQKVALRQEBELKVL 539
Db 481 NAYMAKTKTYQVAHMSKONRIKLANEILNGIKVLYKYAMELAFKQKVALRQEBELKVL 540
Qy 540 KKSAYLSAVGFTFVNCPLVALCTPAVYVITDENNITLDQTAVALFPIILFPNITL 599
Db 541 KKSAYLSAVGFTFVNCPLVALCTPAVYVITDENNITLDQTAVALFPIILFPNITL 600
Qy 600 PMVISTIVQSVSLKRLRIFLSHELEPDSIERRPVDDGGTNSITVYNAFTWARSDDP 659
Db 601 PMVISTIVQSVSLKRLRIFLSHELEPDSIERRSIKSGSG-NSITYKNAFTWARGSDP 659
Qy 660 TLNGITTSIPBGALVAVVGQVCGKSSLSALLAEMDRVBGVAIKGSVAVYPOQAMION 719
Db 660 TLNGITTSIPBGALVAVVGQVCGKSSLSALLAEMDRVBGVAIKGSVAVYPOQAMION 719
Qy 720 DSLARENILFGQQLPEPYRSYVQACALLPDEILIPSDRTEIGKGNLSGGQORVSLA 779
Db 720 DSLARENILFGQQLPEPYRSYVQACALLPDEILIPSDRTEIGKGNLSGGQORVSLA 779
Qy 780 RAYVSNADYLFDDPLSAVDAHVGKHI FENVYIGPKGMLKNKTRILVTHSMSTYLPQVDYI 839
Db 780 RAYVSNADYLFDDPLSAVDAHVGKHI FENVYIGPKGMLKNKTRILVTHSMSTYLPQVDYI 839
Qy 840 VMSGKTSIEMGVSQELLARDGAFAPLRTYASTQEBDAEENGVTVGSGPKAKOMENG 899
Db 840 VMSGKTSIEMGVSQELLARDGAFAPLRTYASTQEBDAEENGVTVGSGPKAKOMENG 896
Qy 900 MLVYDSAGKQLOQDLSSSSSYSQDLSRHNSTALQKAEKERTKLMEDKAOQGVK 959
Db 900 MLVYDSAGKQLOQDLSSSSSYSQDLSRHNSTALQKAEKERTKLMEDKAOQGVK 955
Qy 960 LSYVYMDYKALGFLISFSLFLFCMCHVNSALASNYMLSLMTDD-PIVNGTQEHTKVRLSV 1018
Db 960 LSYVYMDYKALGFLISFSLFLFCMCHVNSALASNYMLSLMTDD-PIVNGTQEHTKVRLSV 1015
Qy 1016 YGALGIIQGAALIFGYSMAVSIIGIIFASRRRLHDLILVYLRSPMSFFERTPSGNLVNRRSK 1075
Db 1016 YGALGIIQGAALIFGYSMAVSIIGIIFASRRRLHDLILVYLRSPMSFFERTPSGNLVNRRSK 1075
Qy 1079 ELDTVDSMIFPVYIKMFMGSLFENVIGACIYIILATPIAAIIPPLGLIYFFVQRFYVASSR 1138
Db 1079 ELDTVDSMIFPVYIKMFMGSLFENVIGACIYIILATPIAAIIPPLGLIYFFVQRFYVASSR 1135
Qy 1139 QKRLSESVSRPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIIVANRW 1198
Db 1136 QKRLSESVSRPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIIVANRW 1195
Qy 1199 LAVRLCEVGNCIIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYANMLVRMSSEMETNI 1258

Db 1196 LAVALECGNCIVLFAALFAVISRHSLSAGLVGSVSLQITAVIMLVMSSEMETNI 1255
Qy 1259 VAVERLKEVSETEKEAPWQIOETAPSSWPPOVGRVEFPNNYCLARRREDLDPVLRHINVTIN 1318
Db 1256 VAVERLKEVSETEKEAPWQIOETAPSTWPSHGRVFPDYCLARRREDLVLKHNINVTIE 1315
Qy 1219 GGEKVGIVGRGAGSSLTLLGFRNESABEGIIIDGINIAKIGLHDLRFKITTIIPODPV 1378
Db 1316 GGEKVGIVGRGAGSSLTLLGFRNESABEGIIIDGINIAKIGLHDLRFKITTIIPODPV 1375
Qy 1379 LFGSSLRNNLDPFGQYSDDEEYVSLLELAHLDKDFVSALPDKLDHEGAGGEMLSVGORQV 1438
Db 1376 LFGSSLRNNLDPFGQYSDDEEYVSLLELAHLDKDFVSALPDKLDHEGAGGEMLSVGORQV 1435
Qy 1439 CLARALLRKTKILVDEXTAAVLETDLLIOSTIRTOEDCTVTIARLNTIMDYTRVI 1498
Db 1436 CLARALLRKTKILVDEXTAAVLETDLLIOSTIRTOEDCTVTIARLNTIMDYTRVI 1495
Qy 1499 VLDRGETOEVGAPSDLLQORGLFYSMADAGLV 1531
Db 1496 VLDRGETOEVGAPSDLLQORGLFYSMADAGLV 1528

RESULT 15
US-10-618-281-63
Sequence 63, Application US/10618281
Publication No. US20040219609A1
GENERAL INFORMATION:
APPLICANT: Day, Anthony G.
APPLICANT: Estelle, David A.
APPLICANT: Lyons, Eric H.
APPLICANT: Yao, Jian
TITLE OF INVENTION: Methods for Modulating Proteins Not
FILE REFERENCE: GC773-2
CURRENT APPLICATION NUMBER: US/10/618,281
CURRENT FILING DATE: 2003-07-11
PRIOR APPLICATION NUMBER: US 60/395,325
PRIOR FILING DATE: 2002-07-12
NUMBER OF SEQ ID NOS: 92
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 63
LENGTH: 1303
TYPE: PRT
ORGANISM: Homo sapiens
US-10-618-281-63

Query Match 61.3%; Score 6068.5; DB 5; Length 1303;
Best Local Similarity 81.4%; Pred. No. 0;
Matches 1229; Conservative 17; Mismatches 57; Indels 207; Gaps 10;

Qy 22 WNTSNPFTKCFQNTVLMVWPCFYLMACFPFYFLYLSRHDGYIOMTPINKTKTALGFL 81
Db 1 WNTSNPFTKCFQNTVLMVWPCFYLMACFPFYFLYLSRHDGYIOMTPINKTKTALGFL 60
Qy 82 WIWCMADLFYSFEMRSRGIFLAPVFLVSPILLGITLLATFLQLERRKVOSSGIMLTF 141
Db 61 WIWCMADLFYSFEMRSRGIFLAPVFLVSPILLGITLLATFLQLERRKVOSSGIMLTF 120
Qy 142 WLVLVLCALALTRKTKTALKEQAQNDLFRDITFYVYFSLLLQLVLSCEFDSPPLSET 201
Db 121 WLVLVLCALALTRKTKTALKEQAQNDLFRDITFYVYFSLLLQLVLSCEFDSPPLSET 177
Qy 202 IHDNPPCRESSASFLSRITTFMMITGLIVRGYRQPLEGSDLMSLNKEDTSEGVVPLVKNW 261
Db 178 IHDNPPCRESSASFLSRITTFMMITGLIVRGYRQPLEGSDLMSLNKEDTSEGVVPLVKNW 237
Qy 262 KKECAKTRKQPKVYVYSKSDPAQPKRESSKVDANEVEALIVYSPOKEWNPGLFVLYKTF 321
Db 238 KKECAKTRN-----SSGSGESCSANTEALF-----PA--PTCHKSF 271
Qy 322 GPYFLMSFPFKAIHDLMMFSGPQILKLIKRVNTRKAPDWGQFYFVTLVFTACLGQTLVL 381

Db 272 QALSL-----LCRLIKRVNTRKAPDWGQFYFVTLVFTACLGQTLVL 314
Qy 382 HQYHIFCVSGMRIKTAVIGAVYRKALVTINSARKSSTVGEIYNLMSVDAQRFMDLATYI 441
Db 315 HQYHIFCVSGMRIKTAVIGAVYRKALVTINSARKSSTVGEIYNLMSVDAQRFMDLATYI 374
Qy 442 NMISAPLOVTLALYILMLNIGPSVLAGVAVMLVMPVNAVMARKTQYQAHKSKDNR 501
Db 375 NMISAPLOVTLALYILMLNIGPSVLAGVAVMLVMPVNAVMARKTQYQAHKSKDNR 416
Qy 502 IKLNEILINGIKVLKYAMELAFKQKYLAIROBELKYLKKSAYLSAVGTFTWVCPFLVA 561
Db 417 IKLNEILINGIKVLKYAMELAFKQKYLAIROBELKYLKKSAYLSAVGTFTWVCPFLVA 476
Qy 562 LCTFAVYVITDENNILDQAFAVSLAFENILRPELNLPMVSIISIVQASVKRLRIPLS 621
Db 477 LCTFAVYVITDENNILDQAFAVSLAFENILRPELNLPMVSIISIVQASVKRLRIPLS 528
Qy 622 HEELFPPSIEERRPYKDGCTNSTVRNATFTWASDDEPTLNGITFESIPEGALVAVGVG 681
Db 529 GATSERPWPWSRPRKHG-----TROASPSVABPVLCRESITFESIPEGALVAVGVG 581
Qy 682 CGKSLSLALLAEMDKVEGHVAIKGSVAVYVPOQAWIONDSLRENILRGCOLLEBPYRSVI 741
Db 582 CGKSLSLALLAEMDKVEGHVAIKGSVAVYVPOQAWIONDSLRENILRGCOLLEBPYRSVI 641
Qy 742 QACALLPDLLEILPSGDRTEIGKGVNLGGQKQKVSARAVSNADYILFPDPLSAVDAN 801
Db 642 QACALLPDLLEILPSGDRTEIGKGVNLGGQKQKVSARAVSNADYILFPDPLSAVDAN 701
Qy 802 VGKHIPEENVIGPKMLKXKTRILVTHSMSTYLPQVDYIIVNSGKISSENGSYQELIARDGA 861
Db 702 VGKHIPEENVIGPKMLKXKSCLI----- 724
Qy 862 FAELRLRYASTEQQAEBENGVTGVSQPKAKOMENGMVLTDSAGKQLOQLSSSSSY 921
Db 725 -----S 725
Qy 922 GDISRHNSSTAELQKAEKKEETKLMADKAQOTGVYLSVYMDYKALGLFISFLSIFL 981
Db 726 CDL-----QYKLSYMDYKALGLFISFLSIFL 753
Qy 982 FMCNHSALASNTWLSLMTDDPIVNGTOEHTKVLSTYVAGLISQGIADVPGYSMAVSIGG 1041
Db 754 FMCNHSALASNTWLSLMTDDPIVNGTOEHTKVLSTYVAGLISQGIADVPGYSMAVSIGG 813
Qy 1042 ILSRCLHVDLLHSILSPSPFERTPSGNLVNRFSEKELDTVSMIPEVIMFPGSLPNV 1101
Db 814 ILSRCLHVDLLHSILSPSPFERTPSGNLVNRFSEKELDTVSMIPEVIMFPGSLPNV 873
Qy 1102 IGACTIVILLATPIAIIIPPLGLYFPVQFYVASSQQLKRLSEVSRSPPVSHNETLLG 1161
Db 874 IGACTIVILLATPIAIIIPPLGLYFPVQFYVASSQQLKRLSEVSRSPPVSHNETLLG 933
Qy 1162 VSVIRAFEBQERFHQSDLVYDENOKAYYSIVANRWLAVALCEVGNCTVLPALPAVIS 1221
Db 934 VSVIRAFEBQERFHQSDLVYDENOKAYYSIVANRWLAVALCEVGNCTVLPALPAVIS 993
Qy 1222 RHSLAGLVGLSVSYSQVTTYLNMLVRMSSEMETNIIVAVBRKLEYSETEKEAPWQIOET 1281
Db 994 RHSLAGLVGLSVSYSQVTTYLNMLVRMSSEMETNIIVAVBRKLEYSETEKEAPWQIOET 1053
Qy 1282 APPSPQVGRVVERFRNCLARRREDLDPVLRHINVTINGEGRVGVIGRTGAGKSSLTGLFL 1341
Db 1054 APPSPQVGRVVERFRNCLARRREDLDPVLRHINVTINGEGRVGVIGRTGAGKSSLTGLFL 1113
Qy 1342 RINESAAGEIITIDGINIAKIGLHDLRFKITTIIPODPVLFGSGLRNNLDPFGQYSDDEEYV 1401
Db 1114 RINESAAGEIITIDGINIAKIGLHDLRFKITTIIPODPVLFGSGLRNNLDPFGQYSDDEEYV 1173
Qy 1402 SLELAHLKDFVSALPDKLDHEGAGGEMLSVGORQVCLARALLRKTKILVDEXTAAVD 1461

Db	1174	SLELAHLKDFVSALPDKLDHECARGENLSVGOROLVCLARALRKTKILVLDATAAVD	1233
QY	1462	LETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVILDKGEIOEYGAPSDLLQORGLF	1521
Db	1234	LETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVILDKGEIOEYGAPSDLLQORGLF	1293
QY	1522	YSMAKDAGLV	1531
Db	1294	YSMAKDAGLV	1303

Search completed: December 15, 2005, 15:46:47
Job time : 169.832 secs

Qy	855	LLAADGAAEAELETRYYASTEGEQDABEENVVTGVSQGRKAEAKOMENGLMTYSAGK--QLOr	912	
Db	360	LLOKRGALVCL-----DQARQPDQR--EETEPGISTKQPRG-----TSAGRPPELR	407	
Qy	913	QLSSSSSYSGGISRHNNSTALOKAEAKEET---WKLMEADRAQOTGOVLLSYVPMYK	968	
Db	408	ERSLKS-----VPEKDRTTSEAQ--TEVPLDDPDAGNAGK--DSIQGRVATYTHLAVLR	460	
Qy	969	ALGLFISFLSIFLMWCNVNSALASNYMLSLMTDDPTVNGTQHTKRLSVYGAIGISQGI	1028	
Db	461	AVGTPLCLYALFLFLCQOVASFRCGYWLSLMADDPVAVGGQOTQALRGIGLGLQAI	520	
Qy	1029	AVFGYSMAVSIIGLILASCLHVDLHSLTRBPMSPFEETPBGNYLNPSPKELTDVDSMP	1088	
Db	521	GLFASMAVLLGGARASRLRFQRLMLDVRBSPISFEETPLGHLINRPSKETDTPVDVIP	580	
Qy	1089	EVIKMFMSLFPNVIGACIVILLATPIAIIIPGLGYFFPQRFYVASSRQKRLSEYSR	1148	
Db	581	DKASLSLMYAAGLLESLVAVATPLAVALLPFLIYAGQSILYVSSCQRLDEASGY	640	
Qy	1149	SPVYSHNETLLGVSVIRAFEEQERFIHQSLKYDENQKAYPSIVANRWLVRLCEVGN	1208	
Db	641	SSVCSHMAETFGSGTVVAFRTQAPFVAQNNAVDSEQRISFPRLVADRMLAAVVELGN	700	
Qy	1209	CIVLFAALFANISRSISAGIVGLSVSLQVTTYLMVLNMSSEMENTIYAVERLXES	1268	
Db	701	GLVFAALAAVSLSKHLSAGLVGFSVSAALQVTLQWVVRNMTDLENSIVSVERMODYA	760	
Qy	1269	ETEKAPAOIQETAPPSMPQVGRREFPNYCLAREDEDPLVRHINVTINGEKVGYGR	1328	
Db	761	WTPKEAPRKLPTCAQAPPMPOGQILEFDFGLRIPELPLAVQGVSFIIHGEKVGIVGR	820	
Qy	1329	TGAGKSSLTLGLFRINESAGEEIIIDGINIAKIGLHDLRFKITIIPODPVLFGSLRML	1388	
Db	821	TGAGKSSLASGLIRQEAABEGIWDGVPILAVHGHLTRKSITIPDPLFPESLRML	880	
Qy	1389	DPFGQYSDEEWTSLIELAHLKDFFSALDKDHECABEGENLVSQORQVCLARALLRKT	1448	
Db	881	DLEQHSDBEAIWAALLETQULKALVASLPGQLQYKCADGEBDLSVQOKOLCLARALLRKT	940	
Qy	1449	KILVDEAETAAVDETDLIOSTRTOPEPDCVLTIFARLNTIMDPTRVIVLDGGEIOEY	1508	
Db	941	QILIIDETAAVDPSTELQOMQLGMSWAQCTVLIARLRSVMDCARVAVLMDKQVAES	1000	
Qy	1509	GAPSDLLQORGLFYSMAKDAGLV	1531	
Db	1001	GSPAGLLAQKGLFYRLAQESGLV	1023	
RESULT 2				
US-11-090-439-24				
; Sequence 24, Application US/11090439				
; Publication No. US20050266442A1				
GENERAL INFORMATION:				
APPLICANT: Squillace, Rachel				
APPLICANT: Weiner, Michael P.				
TITLE OF INVENTION: Immortalized Human Tubercous Sclerosis Null				
FILE REFERENCE: 24318-502				
CURRENT APPLICATION NUMBER: US/11/090,439				
PRIOR FILING DATE: 2005-03-25				
PRIOR APPLICATION NUMBER: 60/556,344				
NUMBER OF SEQ ID NOS: 62				
SOFTWARE: PatentIn Ver. 2.1.1				
SEQ ID NO 24				
LENGTH: 1581				
TYPE: PRT				
ORGANISM: Homo sapiens				
US-11-090-439-24				
Query Match	21.4%;	Score 2115,	DB 7,	Length 1581,
Best Local Similarity	31.9%;	Pred.No. 3e-147;		

	Matches	527;	Conservative	305;	Mismatches	564;	Indels	256;	Gaps	37;
QY	32	CFQNTVLVWPCFYLACFPPEFL	-VLSRHDGYIQMTPLNKTALG----	LLMTVCM	86					
Db	26	CFVDALNVVPHVFLFTFTFFLFI	FGMSQSASKVHHSS-----	TWLFHFGHNRLWILTF	79					
QY	87	ADLVSYSPERSRGIFLAPV-----	FLVSPF-----	LLGITTLATFLIQERRGVQSSG	136					
Db	80	MLLFLVLCGLAEGLSDGVESHHLHYMPGMAFMAVTSV--	YVHNLETSNPKLLI	137						
QY	137	IMTFMWALVCAALAIRSKIMTALKEDAQVDLEPRDITF-----	-VYESLILQVLVS	189						
Db	138	ALVYWTIAFL-----	TKTKIKVKFLDHAIGSQARFCULGTGLVILYIGWLLIVE--	VN 188						
QY	190	CFSDRSPLESETIHDNPPCES-----	SASPLSRITFWITGLYRGYRPLEG	238						
Db	189	VIRVRRYFFETKPREVKR-PEDLODLGRLFQPPVNLISKGT	YMMNNAFITAKHKKYI--	245						
QY	239	SDMSLNKEDISEQVNPVVK--	NMKKECAKTRKQPVKVYVS	SKDPAQPKRESSKVDANE	295					
Db	246	-DLRAIGK-----	LPLAMALTNVQRLC-----	-E	268					
QY	296	EVEALIVSPQ-KEMNSSLFVLYLKTGPEYLMGFFFAKIHDLMMFSGPQIILKLKLV-	353							
Db	269	AFDQAVRDICGTGGAIRALIMWALSHAGRCRLVLSSTR	RIIADLIGFAPGLCITFGIVHLSG	328						
QY	354	--NDTPADWQ-----	-GYFTVLLFVTAQIQTVLV	HOYFHIQVYSGMRK	396					
Db	329	KENDVFGCKTQFLGVYFVSQSEPLANNAVILALFLALLIQRTEPLQASYYVAIETGINLR	388							
QY	397	TAVIGAVYRKA--	LVITNSARKSSTGEIVLMSVDQRFMDLATYINMI	MSAPLOYILA	454					
Db	369	GAIQTKIYNKIMHLS	TNSLMSGENTAQOICNL	VALIDTNQLMWFPEFLCP	NLMAMHVQIIVG	448				
QY	455	LYLLMLNIGPSVLGVA	VMVLMVPMVNA	MMAKTQYQVAAHMSKDNRI	KJLNEELINGHXV	514				
Db	449	VILLYILIGVAILGA	AVIITILAPVQVFAIKLSQAQR	STLEYSNERLKOENMLRGIKL	508					
QY	515	LKUYAMELAFKDYALIRQOEBLKVYLKKSAYLS	AVSGTFTWCTPFLVAL	CTPAVYVTTIDEN	574					
Db	509	LKUYAMENI	PFTRVETTRRKEMTSLRAPAYITSISIFMNTA	IPLAVALITFVGHVSPFKE	568					
QY	575	NILDAQTFVSLAF	NILIRFPLNILPMWISS	IVQASVSLKRLRIFL	SHIEELEPDSI----	630				
Db	569	ADFSVSAFASLS	PHILVLPFLSSVRS	TVKALVSQKLSBFL	SABEREOCAPHE	628				
QY	631	-----	ERRPVK-----	-----	DDGGTV-SITVNAATPTW	653				
Db	629	PTPGSPASKYQAV	PLRVVNRKRPAREBDC	GLTGPLOSLVPSADGDADN	CCVQINGYFTW	688				
QY	654	ARSDPFLNGITFSI	PEGALVAVVGVQVCGKSSLS	SALLAEMDKVEG-----	700					
Db	669	TPDGIPTLSNTIR	IRPRQGLTMVIGVQVCGKSS	LLIALLAGMOKVSGAVF	MSLPDSEIG	748				
QY	701	-----	HVAIKGVAIYVPOQ	AIONDSIRENILEFGCQ	LEEPYRSVIOACLL	747				
Db	749	EDPSERETATD	DLIRKRGPAVAYASQ	KWMLNATVEENII	IESPFNKORYKVI	IEACSLQ	808			
QY	748	PDLLETLSGDS	TEIGEKVNLSSGQOKQVSL	ARAVYSAADYVLPDDPL	ASAYDAVKGKIF	807				
Db	809	PDIDLTPGDQ	TOIGERGINISGGOROKI	SVARALYQCRANVFL	DDPSPALDIHLS	SDILM	868			
QY	808	ENVVIGPKMLK-----	NKTRILVTHSMGYL	PQOVVITYMSGAKIS	SEMOSYOBELARDAF	862				
Db	869	Q-----	AILLELRDRDKT	VVLVTHKLOYLPHAD	MIIMKQGTQIRBETL	ADPQSEQL	923			
QY	863	AEFLRTYASTBO	DAENGVTVGSGP	GEKAKOMENGM	VTDSAGKOLQ	OROLSSSSSYSG	922			
Db	924	FEHWKTLMN-RQDQ	LEKEVLT-----	ERKATEPQD-----	LSRAMSSPDG--	965				
QY	923	DISRHHNSTABLQ	KAKEETWKLMDKA-----	QYGVKLSVYNDYMA	IGLGFISFL	977				
Db	966	-----	LQDEEBEEBAE	EEEDDNLSMLHQAEL	VPWRCATYLS	SAGIILLST	1014			

```

OY 978 SIFLEMCNHSVALSNMYLSMTDPIV-----NGHOET-----KYRLSYGALG 1023
Db 1015 LVFSOLLKHMVLVADIDYWLAKTKDSALTTPTPAANCSSOECTLDQTYVAMFTVLCSLG 1074
OY 1024 ISOGIAYFGVSNVAVSIGGILASRCLVHDLHSILRSPMSFEPTSPGNLVNRFSEKLDTV 1083
Db 1075 I---VLICVTSYVTEWMTGKVKAKRHRSLRNLIILAPRFFETPLGSLINRFSDCNTI 1131
OY 1084 DSMPEVIMKFMGSLFVNIGACTIVLLATPLTAIIIPPLGLIYFVQRFVYVASSRQLKRL 1143
Db 1132 DQHPSTLECJSRSTLTCVSNALAVSYTTPVFLVALTEPLAVICYFIOQKFPVVASRDLQOL 1191
OY 1144 ESVSPSPYSHNETLGVSVYRAFEEOERFLHGDLVADENOKRAYPSIVANPMLAVRL 1203
Db 1192 DDTTOLPLSHFAETVEGLITTRIRRYEARFOQKLEYTDSNNIASLFLTPANRWLEVRM 1251
OY 1204 ECVGNICVLPAALPAVIS--RHSLSAGLVGLSVSYSLQVTTYLNMVLVMSSEMETNIVAV 1261
Db 1252 EYIGACVVLIAAVTSSISNSLHRELSAGLVGLGTALVAVSNYLMNMVNNLADMELQCGAV 1311
OY 1262 ERUKYESTEXE-----APMOIOTETAPSSMPOVGRVYEFNRNYCLIRYEDDLPVLKHNV 1315
Db 1312 KRHHGLKTEABSYEBGLAPSLI-----PKWPPDQGXIOIQLSVRYOSSLKPVLKHYNA 1366
OY 1316 TINGEKYIGVIRTAGKSSLTLLGLFRINESABGEIIIDGINIAKIGHDLRFKTTIIPQ 1375
Db 1367 LISRQKIGICGRISGKSSSFLAFRNVDFEGHIIITDIDIDAKLPHTIRSRSTIILQ 1426
OY 1376 DPLVFSGLRNMLDPFSQSDBEVWTSLELAHLKDQFVSLAPDKLDHCEAGEGENLSVGOR 1435
Db 1427 DPLVFSGITRNLDPERKCSSTLMEALIEIQLKLVAKLQGDALITTEGGENPSSQGR 1486
OY 1436 QLVCLARLLKTKLVLDEBTAAVDELTDLIGSTIRTOEDCTVLTARLNTINDYT 1495
Db 1487 QLFCLARFVFKTSLIFIMDEATASIDMAITEMILQCVWMTAADRVTIVTARVHTIISAD 1546
OY 1496 RVLVLDKGEIOEYGAPSDLLQOR--GLFSPSMK 1526
Db 1547 LVIVLKRAIIIEFKPEKILSRKDSVFASFVR 1578

RESULT 3
US-10-995-561-527
; Sequence 527, Application US/10995561
; Publication No. US20050272054v1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CLO001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 527
; LENGTH: 1588
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-527

```

[illegible]

QY	137	MLTFMFWALCALALASRKTMTLAKEDAOVDLPEDIF-----VYFSLILLOVLS	189
Db	138	ALVYWTIAFL-----TKTKFVKELDHAIGSQLFCLTGLLVIIYGMILLVE--VN	188
QY	190	CFSDRSPLFSSTIHDPNCPES-----SASFLSRTTFMWITGLIVRGYQPIEG	238
Db	189	VIRARRYFFFTPREVKP-PEDLODLGVRLQPFNNLSKGYMMNNAFIKAHKPI--	245
QY	239	SDLSLNKEDTSEOVVPLVK--NMKECAKTRKQPVKVYSSKDPAPRESSKYDANE	295
Db	246	-DLRAIGK-----LPIAMRALNYOGLC-----E	268
QY	296	EVEALIVSPQ-KEMNPFLFKVLYKTFEPYFLMFPFAIDHLMFSGPQILKLIFV-	353
Db	269	AFDQAVRCDIGTQCARAIMQALSHAFRRVLVSTFFIIDLJLQFAGPLCFGIYDHG	328
QY	354	--NDTKAPDWQ-----GYFYTLLEVTACTQTVLHXYFHHCFVSGMRK	396
Db	329	KENDVFGKTOFLGVYFVSSQEPFLANAVALVLEFLALLORTFLOASYYVAIEGIMLR	388
QY	397	TAVIGAVYRKA--LVITNSAKKSTVGEIVNLMVSADAPFMDLAYIMIMSAPLOVILA	454
Db	389	GAIOTKYINXKIMHLSSTNSLWGEWTAQOICNLVALDTQOLMWFELCPNLMAWPOIIVG	448
QY	455	LYLIMLNGPSVLGVAVMVLMVFNANAMAKTKTYQVAHHKSKDNRIKLMNLIANGIKV	514
Db	449	VILLIYLLIGVAGALGAANIILLAPQYVARKLSQAQSTLEYSERLKOJNEMLRGIKL	508
QY	515	LKTIAMELAFKOKVALIROEELKVLKKSAYLSAVGTFTWVCTPFLVALCTPAVYVTTIDEN	574
Db	509	IKTIAWENIEFTRVETTERKEMTSLRAFAIYTSISIFNNTAIPILAAVITFEVGHVSPEKE	568
QY	575	NIILDAQTFVSLAFNILRPFNLTPWYSSIVQASVBLKRLPLSHELEPDSI----	630
Db	569	ADFBPSVAFASLPHILVTLPLFLSSVVRSTKALVSQKLSFSLSSAEIRBECAHRE	628
QY	631	-----ERRPVK-----DGGGTN-SITVRNATFW	653
Db	629	PTPOGPAASKYOAVPLRVNRRKRPAREDCRLTGPLQSLVPSADGADNCCVOIMGQFTW	688
QY	654	ARSDPPLNGTTPSPBEGALVAVVYGQVCGGSSILSALLAMDKYEG-----	700
Db	689	TPDGIPTLSNITIRIPROQLTMIVGQVCGGSSLLAALGMOQKSGAVFWSSLPDSEIG	748
QY	701	-----HVAIKGSVAVYPOQAMIONDSLRNITLGCOLLEBYRYSVIOACALL	747
Db	749	EDSPERETADDLIRKRGPAVAVASQKPMILNAYEENIIFESPNNKRYGMVIEACLO	808
QY	748	PDLBILPSGDRTEIGEKVNLISGQOKORVSLARAVYSNADYIFPDDPLSAVDAYGKIF	807
Db	809	PDIDILPHGDQTOIGERGINLSGQORQISVARALYOHANVFLDDPSALDIHLSHLM	868
QY	808	ENVIGPQGMUK-----NKTRLVYHMSYLPQOVVVIYMSGKISEMSGYOBELLARDQAF	862
Db	869	Q-----AGITELRLDDKRTIVLVTHKHQYLRHADWITLAKKQDTIQRBESTLQFORSEOL	923
QY	863	AEPFLATYASTBOADAENGVLTVGSGPKEAKOMENGMVLVDSAGKOLQOROLSSSSSYSG	922
Db	924	FEHMKILMN-RQODELEKETV-----ERKATPEPOG-----LSRAMSDGL--	965
QY	923	DISHHNSTAELQAEAKKETWKLMBDKA-----QTGVKLSVYNDYMAALIGLISFL	977
Db	966	-----LODEEBEEBAASEBDNLSMLHQALETWRCACATYLSAGILLSTL	1014
QY	978	SIFLPMGNHVSALASNWVLSMTDDPIV-----NGTOEHT-----KRLSYGALG	1023
Db	1015	LVFSQOLKHMVLAVIDWYLAKTDSALTLTPAANCNLSQGCITLDQTYAVAVPVLCSIG	1074
QY	1024	ISGCIANFGVSMASVIGIILASRCLHVDLHLSIRSPMSFPERIPSGNLVNRFSKELDTV	1083
Db	1075	I---VLCIVSVTEWMTGLKVAKKLSHLSNRIITLAPRRFETTPJLSSINRFPSSDCWTI	1139

Db	966	-----L00BEBEEBAE5EDDNLISMLHQRAETPMTACAKYLSAGILLSL	1014
Qy	978	SIFPMCNHVSALASNYWLSIMTDPYV-----NGTQENT-----KRLSYGALG	1023
Db	1015	LVFSQLLKHMYLVAIDYWLAKMTDSALTLPAPRANCSSLQGCITLDQYVAVVFLYLSIG	1074
Qy	1024	ISGCIANFGYMAVSIIGITLASRCLHYDDLHSILRSPMSFFEPKPSGULVNRFSKELDTV	1083
Db	1075	I---VLCIVTSVTEWMTGLKVAKLHSLNLNRIITLAPRPFETPLGSLINRFFSDCMTI	1133

Qy	1084	USMTPEVIMKMGSLFENVIGACIVYLLATPTLAAITIIPELGLIFPVQGFYVASSRQLKRL	1143
Db	1132	DOHLPSTIECSRSTLLCVSALAAVSYVLPFVLLALVLAIVCYFIQKRFVVASDLOOL	1191
Qy	1144	ESVSRSPYSHFNFTLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIYANMELAVRL	1203
Db	1192	DDTLPLPLSHFAEFVEGLTIIIRAFRYEARFQOKLBYTDSNNIASLPLTMANRLLEVRL	1251
Qy	1204	-----ECVGNCLYLPALFAVIS--RHSLSAGLVGVSYSLOVTTYLLMWLRMSEM	1254
Db	1252	ATPLPEQEIYACAVLLAAVTISINSLSHRELISAGLVGTLVYALMVSNYLMMVRYNLADM	1311
Qy	1255	ETNIVAVRLKEVSETEK-----APMQIDETAPSSMPQVGRFERNYCIARVEDLDF	1308
Db	1312	ELQIGAVKRIHGLKLTENESYBGLLAPSLI-----PKMPDQKQIQIOWLSVRDYSSLKP	1366
Qy	1309	VLRIHNTVINGEKVGIIVRTGACKSSLTGLFRINSAEGEIIIDGINAKIGLHDURF	1368
Db	1367	VLKHNALISPGQKIGICGRTSGGKSSFLAFFRMYDPFEGHIIIDGIDIAKLP.LHTRS	1428
Qy	1369	KITIIIPQDPV.FSGSLRMNLDPFSGYSBEVWTSLELHKLDFVSLDYLKDHCEACGGE	1428
Db	1427	RLSIILOPVLFSGIRFNLDPERKCSSTLWEALIEIQLKLVVALPGGLDAILTBGGE	1486
Qy	1429	NL.SVGQROLVCILARALLKRTKILVLDERTAAVDETDDLIOSTIRQPEDCTVLTIAHRL	1488
Db	1487	NFSQOQRQLFCIARAFVAKTISIFIMDEAPASIDAMTENILQCVMTAVPADRVVTIARV	1546
Qy	1489	NTIMDYTRYVILVDKEIOEYCAPSLLQOR--GLFYSMAK	1526
Db	1547	HTILSADLVILVKRGALLFPDPEKELLSRKOSVFSFVR	1585

```

RESULT 4
US-11-090-439-26
: Sequence 26, Application US/11090439
: Publication No. US20050266442A1
: GENERAL INFORMATION:
: APPLICANT: Squillace, Rachel
: APPLICANT: Weiner, Michael P.
: TITLE OF INVENTION: Immortalized Human Tuberosus Sclerosis Null
: TITLE OF INVENTION: Anionomyolipoma Cell and Method of Use Thereof
: FILE REFERENCE: 29318-502
: CURRENT APPLICATION NUMBER: US/11/090,439
: CURRENT FILING DATE: 2005-03-25
: PRIOR APPLICATION NUMBER: 60/556,344
: PRIOR FILING DATE: 2004-03-25
: NUMBER OF SEQ ID NOS: 62
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 26
: LENGTH: 1581
: TYPE: FRT
: ORGANISM: Homo sapiens
US-11-090-439-26

```

Query Match	21.2%;	Score 2100;	DB 7;	Length 1581;
Best Local Similarity	31.7%;	Pred. No. 3.7e-146;		
Matches 524;	Conservative 306;	Mismatches 567;	Indels 254;	Gaps 377;

[illegible]

Dd	189	VIRXRRIIFKFTREPVKRPEDLDLOLGRFLOPFPNLPSKTYMMMAFITAHKKPI---	245
Qy	240	DLWSLNKEDTSEQVNPVLVK---NMKCEAKTRKQPKVVYSSKDPAPQYESSKVDANEE	296
Dd	246	DLRAIGK-----LPIYMRALTYQRLC-----	269
Qy	297	VEALIVSPQ-KEMNPBLFKVLTYTPEBPYFLMSPFFKAHDLMMGSPQILKILIKRV--	353
Dd	270	FDAQVRDIOQTQGARAIWQALSHAFRRRLVLSFTFILLDLGFPAPLCIFGVLDHLK	329
Qy	354	-NTRKAPDMQ-----GYFYVLFPVTACLOTLYLHQFHCIFVSGMRIKT	397
Dd	330	ENDVPQKRTQPLGYFPYSSQEFLANAVLAVLFLFALLLOQRTPLQASYYVALIEGIMLRG	389
Qy	398	AVIGAVYRKA--LVITNSARKSSITGEIYNLMSVDAQRFMDLATYINMWSAPLOVILAL	455
Dd	390	AIQRIKYNKIMHLSNMLSMGEMTAGQICNLVADTNQLMWFPLCPNLMMAPQOITGVG	449
Qy	456	YLLMLNIGPSTLAGVAMVLMVNPVNAVMANKTKTYQVAHKKSKDNRKLMNELLNGIKVL	515
Dd	450	ILLYYIIGVSLISGAAYIILLAPQYVARKLSQAQSTLEYSNERLKQNMELRGIKLL	509
Qy	516	KLYAMELAFKQVLAISOEBLKVYKCAYSASVATFPWVCPLFVALCTFAVYTYIDENN	575
Dd	510	KLYAMENIFRRVYVTRKREKTSIRAPAIYTSISIFMWTALPIAAVLITFGVHVSFFKEA	569
Qy	576	ILDAQTAFFVSLATFNILRFPNLILPMYISSIVQASVSLKRLRIFLSHEBELPDSI----	630
Dd	570	DFSPSAFASLSLPHILVTPLFLLSSVSRSTVKALVSQKLSEFLSASAIREEQCAHPER	629
Qy	631	-----ERRPVK-----DGGTN-SITVRNATFTWA	654
Dd	630	TPQGPASKYQAVPLRVNRRKPAEDCRGLTGPLQSLVPSADGADNCCVQIMGTYFTWT	689
Qy	655	RSDEPTINGIFESIPEGALVAVGVQVCGGKSSLLSALLAEMDKYEG-----	700
Dd	690	PDGIFPILSNITIRLPRGOLTWIVQVCGGSSLLALLGEMQKXGAVFMSSLPDSERGE	749
Qy	701	-----HVAIKGSVAYVPOQAMIONDSLRENILFGCQDEBPYRSVIGACALLP	748
Dd	750	DPSPERETANDLIRKKGPAVAYASQKXMLNATAENIIFESPNNQRYKMWIEACSLQP	809
Qy	749	DLEILPSSDRTEIGEKVNISSGGQKORVSLARAVYSNADYILFPDPLSANDAHYKHIFE	808
Dd	810	DIDILPMDQOQI GERGINISSGGORRISVARALYOHANVFLDDPPSALDIHLSDHMQ	869
Qy	809	NVIGPKGMLK-----NKRILYTHSMYSLPOVDYIYMSGGKISEMSSYOELLARDGAFA	863
Dd	870	-----AGLEILRDDKDKTIVLVTHKLOVLPRADWTIMKOSTTOREBGLTQFORSBQCLF	924
Qy	864	EFLRTYASTEOQDAEENGVTGVSQPEKAQKQMNGLMTDSAGKOLOROLSSSSSYSGD	923
Dd	925	EHWKTIWN-RDQOLEKETVY-----ERKATEPQG-----LSRAMSSSDGL---	965
Qy	924	ISRHHNSTALQKAEAKKETWKLMEADKA-----QTGQVKLSYUWDMYKALIGLFISFLS	978
Dd	966	-----LQDEEBEEBAEASEEDNLSMLHQAELIPWRACAKYLSAGAILLSIL	1015
Qy	979	IFLPMCHVSLASNYMLSLMTDDPIV-----NGQEHIT-----KRLSYYGALGI	1024
Dd	1016	VFSQLLKRMVLAIDWYLAKTWDSALITTPAARNCSSQECTLODTQYVAMFYVLCSTGI	1075
Qy	1025	SQGIATVFGYSAVSIIGLILARCSLHDVLDHSILRSPMSFFERTSPGNUMVNSFKELPTVD	1084
Dd	1076	---VLCIVTSTYVEMTGKXAKRHLRSLNRIILAPKRFEEFTTLLGSLINRFSDDCNTID	1132
Qy	1085	SMIPEVIMKMGSLFNVIACIVILLATPIAAIIIPGLGIYFVQRPYVASFQKRL	1144
Dd	1133	QHIPTECTLSRSTLLCVSALAVISYVTPVLAIVALLPLAVVCYFQKRFVARSDDLQOLD	1192
Qy	1145	SVSRSPVYSHNETLLGVSIVTARAEQGRFHQSDKVDENQAKAYBSIVANRLAVLLE	1204
Dd	1193	DTTQPLFLSHAEVVEGJTTIRARFYERFQOQKLEVDYDNNISLFLTANRMLVEVME	1252

Db 181 LLRKTQILLDEATAVDPGTELOQWAMLSGWSFAQCTVLLIAHRLSVMDCAVLMVDK 240
QY 1504 EIOEYGA PSDLLQORGLFYSMAKXAGLV 1531
Db 241 QVBSGSPAQLLAQGLFYRLAQSGLV 268

RESULT 9
US-10-613-744-13

; Sequence 13, Application US/10613744
; Publication No. US20050272093A1
; GENERAL INFORMATION:
; APPLICANT: MacKinnon, Roderick
; TITLE OF INVENTION: The Rockefeller University
; TITLE OF INVENTION: Assays for Screening Compounds which Interact With
; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
; FILE REFERENCE: 018512-002901US
; CURRENT APPLICATION NUMBER: US/10/613,744
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US/09/275,252
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: US 09/045,529
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/054,347
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: WO PCT/US99/06307
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-613-744-13

Query Match 7.9%; Score 786.5; DB 6; Length 391;
Best Local Similarity 46.2%; Pred. No. 1.8e-50;

Matches 150; Conservative 75; Mismatches 89; Indels 11; Gaps 6;

QY 1564 RTERRARFVSKKNCNVAAHKNIREQGRFL--QDFTLVLDLKPHTLLIFTSPLCSWL 1621
Db 35 RSRQ-RARLVSKERKCNIEFGNVDAQSRFIFVDIMTTVLDLKKRYKMTVITRFLGSWF 93
QY 1622 LFAMVWMLIAFAHGDL---APGEGTNPCTSTSHSSAFLEIEVQVITGFGGRMYTE 1677
Db 94 LFGLLWTVVAVVHDLPEFYPPDNR--PCVENINGMTSAFLFLETOVITGYGFRFVTE 151
QY 1678 ECPALILILYQNTVGMINAIMLGCIFMKTQAHRRAETLIFSKHAVITLHGRLEML 1737
Db 152 QCATATLFLFQSLIGVINSFWCGALIAKISRKKRKTITFSKNAVISRGGKCLLI 211
QY 1738 RVGLDKRSMITATIHQVVRKTSPEGEVVPRLHQVDIPMENGVGNGIPLVAPLITYH 1797
Db 212 RVNLRKSLIGSHIYKGLKLTITTPGETIITLDQTNINFWVDGNGNLFISPLITYH 271
QY 1798 IDSNSPLYDLAPSDLHHQDLEIIVLEGVETTGITTQARTSYLADEILMGORFPIVA 1857
Db 272 IDHNSPFFHMAAETL--SQODFELVFLDGVETSTATCQVTSYVPEEVLMGYRFVIVS 330
QY 1858 E-EDGRSYVDYSKRGNTIKVPTPLC 1881
Db 331 KTKGKRYVDFHNGKTVETVETPHC 355

RESULT 10

US-10-793-626-326
; Sequence 326, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US

CURRENT APPLICATION NUMBER: US/10/793,626

CURRENT FILING DATE: 2004-03-04

PRIOR APPLICATION NUMBER: 60/164,258

PRIOR FILING DATE: 1999-11-09

NUMBER OF SEQ ID NOS: 4472

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 326

LENGTH: 548

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: synthetic

US-10-793-626-326

Query Match 5.2%; Score 514.5; DB 6; Length 548;
Best Local Similarity 27.1%; Pred. No. 3e-30;

Matches 157; Conservative 110; Mismatches 211; Indels 101; Gaps 16;

QY 1002 DRYNGQEHTRKRLSYGALGISQGIYAVGYSMVAISIGILASRCLHVDLHLSLSPM 1061
Db 12 DGVIN--NHSLTNQEKFSHGVAGIALF-----IFLIVRPI 47
QY 1062 SFPE---RTPSGNLVNRFSKEL-----DTVDSMPIEYIK-----MEMG 1096
Db 48 EFTIQYLAQMTSKILDIRKQLYNHLQALSVRYANNQGVLSRVINDVEQTKDPLIT 107
QY 1097 SLFNVIGACIVILLATPI-----AAIIPPLGL--IYFVQRFYVASSRQLKLE 1144
Db 108 GLNMIMWDCITITIIALSIMFPLDVKLPFAAIFIPPYILVYFFPGL-----RKLTVR 162
QY 1145 SVSRSPYSHPNETLAVSVITRAFEQERFIHOSDLKVDENOKAYYSIAANRLAVRLE 1204
Db 163 SQALAEVQGLHERVQGSVYKSPALIEDNAKPD--NNHKVFLQRAFOHTRMNAYSFA 219
QY 1205 CV-----GNCIVFALFVVISRHSLSAGLVSVSYSLQVTTYLMVIMVMSSEMTNI 1258
Db 220 AINTVDTLGLPIYVGVSYLAIT-GSTTVGTIAFVYLEQRLRLVSSFTTLTQSF 278
QY 1259 VAVERLKEVSETEKE-----APWQIQETAPSSWPQVGRVERFNNCLRYREDLDFVR 1311
Db 279 ASMDRVQLMDEVDYDINGIGAQPIKSK-----QIDLKHSVFNKNEKEVYLH 328
QY 1312 HINVTNGEKVGVGTGQKSLTIGLPRINESAGEIITIDCINIAKGLMDLRKIT 1371
Db 329 DINTLTNGKGTVAFEVSGGSKSTLNLIRFYDVTOGETLIDHNKVCDFLTGSLRNQIG 388
QY 1372 IIPQDPVLPFSGSLRMNL---DPSQYSEDEWTSLELAHKDPFSAIPDKLDHRCARGE 1428
Db 389 LVQODNITLFSDYAKENLILGRP--DATDDEVBAKKAANAHDFISLNPNGYDTEVGERGV 446
QY 1429 NLSVGRQLVCLARALRKTILVLEATAVAVDLETDDLIOSTIRTOFEDCTVLTIAHL 1488
Db 447 KLSGGQKQRISTARIFLNNPVLIDBATGALDESALIIQELMDVLSTKORTLIIAHL 506
QY 1489 NTIMDYTRVIVLDKGEIQEYGA PSDLLQORGLF---YSM 1524
Db 507 STTHADRIYVMENGRIVETGTHQQLINKGAYEHLYSI 545

RESULT 11

US-10-467-657-1346
; Sequence 1346, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657

```
/ CURRENT FILING DATE: 2003-08-11
/ PRIOR APPLICATION NUMBER: GB-0103424.8
/ PRIOR FILING DATE: 2001-02-12
/ NUMBER OF SEQ ID NOS: 9218
/ SOFTWARE: SeqMan9, version 1.04
/ SEQ ID NO 1346
/ LENGTH: 622
/ TYPE: PRT
/ ORGANISM: Neisseria gonorrhoeae
US-10-667-1346

Query Match      5.0%; Score 495.5; DB 6; Length 622;
Best Local Similarity 28.7%; Pred. No. 9,1e-29;
Matches 154; Conservative 106; Mismatches 213; Indels 63; Gaps 13;

QY 1022 LGISGIVFGYMAVSIIGGLASRCLHVDLHSLRSPMSFPERTPSGNLVNRFSKELD 1081
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 108 LVIRGICRFSTYLTMTWVSWTISKIKDMFARKMLTSSRYHQTPSGYVLMNMLNLT 167
QY 1082 TVDSMIPVIMFGSLFNVIGACIVLL---ATPIAIIIPPLGLYFVQRFYVASS 1137
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 168 QSVSNASDIFVTLFRDWTIVTGLTVLLYLMWQSLIVLMFPLSL---SRY---R 220
QY 1138 RQLKRLBSVSRSPVYSHN---ETLLGYSVIRAFEGORFTHOSPL-----KXDE 1184
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 221 DRLKHVSDSQSKSIGTMNVIAETHQGRHVYGLFNGQAQANRFAVNRITVRLSKKITQ 280
QY 1185 NOKAYYP--SIVANRWLAVRLECVNGCIYLFPAALPAVTSRHSLSAGVLGSYSLOYTT 1242
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 281 ATAAHSPSELIALALAV-----VIFALMQSONGYTTIGEPMAFIVA-MLOWYA 330
QY 1243 YLNLVWRSSSEMETNIVAVERLKEYSETEKE-----APWQIQETAPSSMPQVGRVEFR 1296
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 331 PIKSLANISIMQWTFLAADGVCALDTPPEODKGTLPQRVE-----GRISFR 379
QY 1297 NYCARYRDDLDPVLRHINVTINGEKVIGRTGAGKSLTILGPRINESAGEIITIGI 1356
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 380 NVDAVEYRSDGIRKALDNFMIDIRQGERVALVGRSGSGKSTVNLIRFVEPSGNCIDGI 439
QY 1357 NIAKIGLDLRFKTIITIPQDPVLFSGSLRMNLDPFSQYS-----DEWYTSLEIAHKDF 1411
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 440 DIADIKDCLARQPAVLSQDVFLLFDITLPEVY---RISRPAGAEVLSALQANLDSL 495
QY 1412 VSALPDKLDHCEAGGENLSVGOQRLVCLARALLRKTKILVDEATAVADLETDDLIOST 1471
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 496 IDASPLGLHQPISGNSMLSGGQRQVVALARAILADAPILLDEATVSLDNSESRLVQA 555
QY 1472 IRTQPEDCTVLTIAHRLNTIMDYTRVIVLDKGEIQEYGAPSDLLQORGLFYSMAOD 1527
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 556 LERLMENRTGIIVAHRLTTVESADRIIYWDGKKIIEGTHDQLMFQNG--YVTMLRN 610

RESULT 12
US-11-082-389-148
/ Sequence 148, Application US/11082389
/ Publication No. US20050244935A1
/ GENERAL INFORMATION:
/ APPLICANT: Pompeju, Markus
/ APPLICANT: Krogger, Burkhard
/ APPLICANT: Schröder, Hartwig
/ APPLICANT: Zeider, Oskar
/ APPLICANT: Haberhauser, Gregor
/ TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
/ TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
/ TITLE OF INVENTION: TRANSPORT
/ FILE REFERENCE: BGI-131PCN
/ CURRENT APPLICATION NUMBER: US/11/082,389
/ PRIOR FILING DATE: 2005-03-16
/ PRIOR APPLICATION NUMBER: US 09/603024
/ PRIOR FILING DATE: 2000-06-23
/ PRIOR APPLICATION NUMBER: US 60/141031
/ PRIOR FILING DATE: 1999-06-25
/ PRIOR APPLICATION NUMBER: US 60/143262
```

```
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: US 60/151281
/ PRIOR FILING DATE: 1999-08-27
/ PRIOR APPLICATION NUMBER: DE 19930487.4
/ PRIOR FILING DATE: 1999-07-01
/ PRIOR APPLICATION NUMBER: DE 19930489.0
/ PRIOR FILING DATE: 1999-07-01
/ PRIOR APPLICATION NUMBER: DE 19931549.3
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19931550.7
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19932134.5
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: DE 19941379.7
/ PRIOR FILING DATE: 1999-08-31
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 446
/ SEQ ID NO 148
/ LENGTH: 549
/ TYPE: PRT
/ ORGANISM: Corynebacterium glutamicum
US-11-082-389-148

Query Match      4.7%; Score 468.5; DB 7; Length 549;
Best Local Similarity 26.9%; Pred. No. 7,3e-27;
Matches 141; Conservative 110; Mismatches 201; Indels 73; Gaps 18;

QY 1031 FGYSMAVSIIGGILASRCLHVDLHSLRSPMSFPER-----TPSGNLVNRFSKELD 1084
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 50 FGSKLSMRVGR-----DLRSALRGKVNFESEREMGQGAAS--LITNTNDVQOQ 98
QY 1085 SMIEPVIKMFQSLFNVIGACI-----VILATPIAIIIPPLGLYFVQRFY 1133
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 99 MLVQWSTSLMSAPMLAIGIIMAVRODLGSLMWLVISIPVLITV---ALIIYRVDEL 155
QY 1134 VASSRQLKRLBSVSRSPVYSHNFTLGVYIRAF---EEOEPRIHOS-----DLKVD 1183
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 156 QTMOKRIDRIHQIIR-----SOLTGIRVIRAFVREDVERERTTASKVADIGVRTG 207
QY 1184 ENOKAYYPYSIYANRWLAVRLECVNGCIYLFPAALPAVTSRHSLSAGVLGSYSLOYTT 1243
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 208 NLMALMPFAVY---LIMNLASA---YIVGCAQVSEGET-QIGTLFAPQIYIMQILMG 259
QY 1244 LNLVWRSSSEMETNIVAVERLKEYSETEKEAPWQIQET-APPSSMPQVGRVEFRNYCLRY 1302
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 260 VMAAFVFWMPVRAVSAADRIGEVLETTPSY--QAPETPAQPS--SAGEIVFNNAFPAY 315
QY 1303 REDDLPVLRHINVTINGEKVIGRTGAGKSLTILGLP-RINESAGEIITIGINIAKI 1361
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 316 PGADDPVLTNNVSFRVAPSTTAIIIGSTSGKTTL-IGLVPRLFVTEGDVVDGTVDEF 374
QY 1362 GLHDLRFKTIITIPQDPVLFSGSLRMNLDPFSQYSD-----EWMYTSLEIAHKDFVSLAP 1416
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 375 EPLKLMDRIGLVPOKSLFSGTIASNL---RYGNEDATETQLOALIAQADPVREMP 430
QY 1417 DKLHCEAGGENLSVGOQRLVCLARALLRKTKILVDEATAVADLETDDLIOSTIRTOF 1476
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 431 EGLDSEIAGGNNVSGGQRRLAARALLKQPEIYIPDSDSALDVSTDALRRALSTNL 490
QY 1477 EDCYVLTIAHRLNTIMDYTRVIVLDKGEIQEYGAPSDLLQORGLF 1521
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 491 PDATGLIIVAGVSTIRDAQIVVDVNGEIVGITHNTLNTCGTV 535

RESULT 13
US-10-995-561-575
/ Sequence 575, Application US/10995561
/ Publication No. US20050272054A1
/ GENERAL INFORMATION:
/ APPLICANT: CARGILL, Michele et al.
/ TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
/ TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
/ TITLE OF INVENTION: DETECTION AND USES THEREOF
```

```

: FILE REFERENCE: CLO01559
: CURRENT APPLICATION NUMBER: US/10/995,561
: CURRENT FILING DATE: 2004-11-24
: NUMBER OF SEQ ID NOS: 85702
: SOFTWARE: FastSeq for Windows Version 4.0.
: SEQ ID NO 575
: LENGTH: 701
: TYPE: prt
: ORGANISM: Homo sapiens
:
US-10-995-561-575

```

Query Match	4.2%;	Score 419.5;	DB 6;	Length 701;
Best Local Similarity	25.2%;	Pred. No. 4.3e-23;		
Matches 149;	Conservative 130;	Mismatches 234;	Indels 79;	Gaps 21;

```

QY      977  LSIPLFM--CNHVASLASNMYLMSLTDDPVIYNGOEHKRVLSYVAGLIGIOGLAVPEYS 1033
Db      140  LSLPLVLVLTSSLGEMALPFFTGRLTILQDQSGADTTRNLTMSILTLTASAVLEF--- 196
QY      1035 MAVSIG-GIIASRCILV-----DLHSLILSPMSFFERTSGNLVNFSEKELDVDSMI 1087
Db      197  ----VGDSIYNNVTMGHVSHLQGEVFGAVLQETTFPOQNOQTGNMSVTEBTDITLSLSL 252
QY      1088 PEYIKMEGSLFENYIGACTV-ILL-----ATPIAIIIPPLGIYFVQRFYVASSROLK 1141
Db      253  SENLSLFLMTYL--VRGLCLLIGIMMGSVLTMVTLITPLFLPLPKKGVKXVQLLEVOVR 310
QY      1142 RLEVSRSPPVSHNENELLGSVIRAFEEORFIHQSLKXDE---NOK---AYPSIV 1194
Db      311  --ESLAKSQA--TEALSAMPYVRSFANEGEAKREFKEIQETKTLNOKKAVY---A 362
QY      1195 ANRW-----LAVRLEECVNCNTYLFALPAVISRHSLSAGVLGSYSLSQVTTYLNM 1246
Db      363  VNSMTTISIGMLKLVGILYIGQLVTSQ-----VSSGNLVTFYLMQFQVAVEV 413
QY      1247 LVNRSSEMETNIVAVERLKEYSETEKEAPWOIQETAPPS---WPQYRVEFRNYCLRY 1302
Db      414  LLSIYPRQKAVSGSEKIFEYLDRTPR-----CPSPGLTLPRLHEGLVQFQVSPAY 465
QY      1303 --REDDLRFVLPHINVTINGCEKVGIVERTGACKSLSLTGLRINESAGEELIIDGINLAK 1366
Db      466  PNRRDV--LVLOGLFTFLTPRGSEVTALVGNPGSGKSTVAALLONLVQPTGGOLLDBGKPLPQ 524
QY      1361 IGLHDLRFKTIIPQDPVLFGSGLRMLN--DPFSQSPDEEYVTSLELAHLKDPYSALPDK 1418
Db      525  YEHYLYLRQVAAVGOEPOVFRGSRQDENIAYGLTQKPTMEETLTAAYVSGAHSFISGLPQG 584
QY      1419 LDHECAEGENL SVGORQVCLARALFKTKILVLEDEATAVD---LETDDLQSTIRT 1474
Db      585  YDTEVDEBAGSQLSGGROQAVALABALTRKPCVLLIDPATSLADNNSQVQVGLLYES--P 642
QY      1475 QFEDCTYLTTHNRJLNTIMDYTRVVLVDKGELOEFGABSDLIQOGLFPTSMXK 1526
Db      643  EYRSRSVLLTQHLSTLVEQADHILFLREGAIREGTHQOLMEKKGCVYAMVO 694

RESULT 14
US-10-995-561-574
; Sequence 574, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 574
; LENGTH: 808
; TYPE: PR1

```

ORGANISM: Homo sapiens
US-10-995-561-574

Query Match	4.2%;	Score 419.5;	DB 6;	Length 808;
Best Local Similarity	25.2%;	Pred. No. 5.3e-23;		
Matches 149;	Conservative 130;	Mismatches 234;	Indels 79;	Gaps 21

QY 977 LSIFLEM--CNHVSALASNYWLSLWTDPIVNGTQEHKVRLSVYGALGISQGIAVFGYS 1034

```

Db 247 LSLFLVLVWSSLSGEMALPEFTGRITLTDITLDDGASDITTRILITLMSILITISAVLEF--- 303
Qy 1035 MAVSIG-GILASRCILHV-----DLHSILRSPMSFPERTESGVLNVRFSKELDTVDSMI 1087
Db 304 ----VGDSIYNNTMGHVASHLQGBFAGVAVLQETFEFPQONQGNMRSVTEBTDITSLBSL 359
Qy 1088 PEVIKFMGSLFNVIAGIV-ILL-----AFPIAIIIPGLIYFVQRYVAVSSROK 1141
Db 360 SENISLFLMWY--VAGLCILGIMLGSVILMVTILITPLFLPKKGVKXQVLEVEQVR 417
Qy 1142 RLEVSRSRPVYSHNETLLGSVIRAFEEORFIHQSLKXDE---NOK---AYPSIV 1199
Db 418 --ESLAKSQVA--IEALSAMPYVASFANERBEAQKREKQOEIKTLNQKAVAY----A 469
Qy 1195 ANRW-----LAVRLCEVGCNICYLPAALFAVIRSHSISLGVLSVYSLOVTTYLNM 1246
Db 470 VNSWTTISGMLKXGCIYIGGOLVTSQ-----VSSGNLTFVLYQVQTOAVEV 520
Qy 1247 LVRRSMSEMETNIIVAVERLKEYSETKEAPWOIQETAPRSS-----WPQVRVEFRNYCLRY 1302
Db 521 LLSIYPRQKXAVGSEKIFEYILDRPP-----CPFSGLITPLHLEGLVQFOVQVSPAY 572
Qy 1303 --REBDFVLRIHINVTJINGEKEVIGRTGAKSSLTGLFRINSAEGEIIIDGINIAK 1366
Db 573 FNRPDV-LVLOGLFTLARPGEVITALVGNNGSEKSVAAALLQNLVQPTGQGLLIDOEKPLPQ 631
Qy 1361 IGLHDLRFKTIITIPDDPVLFSGSLRMT--DPFSQYSDBEWTSLSELALHKLQFVASLPDK 1418
Db 632 YEHRILRHQVAVAGCEPQVFGRSIOENIAVGLTQPRMEETITAAVAKGASHFISGLPQG 691
Qy 1419 LDHECAEGENLSVQOROLVCLARLAKTKTIVLDEMTAVD----LETDDLQSTIRT 1474
Db 692 YDTEVDGSGQLSGQORQAVALARLIRKPCVLLIDDATSALDANSQLOVQSLYAES--P 749
Qy 1475 QFEDCTVTLIARLNTIMDYTRVILVLDGEIOEYAGSDGLDQOGRFLPSMAK 1526
Db 750 ERYRSVLLIIOHLSVQADHITLFBGCAIRBGGTHQOOLWEKKCCYVAAWO 801

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:27:22 / Search time 176.366 Seconds
(without alignments)
4800.712 Million cell updates/sec

Title: US-10-665-283-1
Perfect score: 9903
Sequence: 1 MALRGFCSDADGSDPLMDMNV.....RSVAVAKXKPFISIPDLSL 1927

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_21:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*
- 9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9903	100.0	1927	9	ADY86938 Human MRP
2	9891	99.9	1927	9	ADY86943 Human MRP
3	9873	99.7	1947	9	ADY86945 Human MRP
4	9734	98.3	1891	2	AAW57486 Human MRP
5	7860	79.4	1531	2	AAW74471 Human mul
6	7860	79.4	1531	2	AAW98994 Human mul
7	7860	79.4	1531	3	AAW55799 Human mul
8	7860	79.4	1531	3	AAW78873 Human mul
9	7860	79.4	1531	5	ABG61810 Prostate
10	7860	79.4	1531	6	ADB20865 MRP1 base
11	7860	79.4	1531	6	ADB20865 MRP1 base
12	7860	79.4	1531	7	ADB87954 Human UGT
13	7860	79.4	1531	7	ADB87954 Human UGT
14	7860	79.4	1531	7	ADB92128 Human MDR
15	7860	79.4	1531	7	ADB92128 Human MDR
16	7860	79.4	1531	7	ADB92128 Human MDR
17	7860	79.4	1531	8	ADN97111 MRP1 prot
18	7860	79.4	1531	8	ADN97111 MRP1 prot
19	7860	79.4	1531	8	ADN97111 MRP1 prot
20	7856	79.3	1530	2	ADN95929 Human BCC
21	7849	79.3	1531	2	AAW57487 Human mul
22	7849	79.3	1531	2	AAW57487 Human mul
23	7849	79.3	1531	2	AAW57487 Human mul
24	7849	79.3	1531	2	AAW74470 Human mul

25	7849	79.3	1531	2	AAW98993 Human mul
26	7849	79.3	1531	3	AAW55798 Human mul
27	7849	79.3	1531	3	AAW78872 Human mul
28	7849	79.3	1531	3	AAW95882 Human mul
29	7763	78.4	1515	8	ADU24094 Human mul
30	7610	76.8	1489	2	AAW96952 Human mul
31	7587.5	76.6	1482	8	ADN03902 Antipsoi
32	7587.5	76.6	1482	8	ADP23422 PRO polyp
33	7572	76.5	1481	5	ABP52108 Homo sapi
34	7560.5	76.3	1482	2	AAW96953 Human mul
35	7535	76.1	1475	9	ADX06206 Human mul
36	7510.5	75.8	1472	9	ADX06204 Cyclin-de
37	7500.5	75.7	1472	2	AAW96954 Human mul
38	7230.5	73.0	1530	7	ADP56611 Bovine MR
39	7195.5	72.7	1416	9	ADX06208 Human mul
40	7194	72.6	1417	2	AAW96955 Human mul
41	7037	71.1	1388	7	ADJ69912 Human hea
42	7002.5	70.7	1528	2	AAW57487 Human mul
43	7002.5	70.7	1528	2	AAW74472 Mouse mul
44	7002.5	70.7	1528	2	AAW98995 Mouse mul
45	7002.5	70.7	1528	3	AAW55800 Murine mu

ALIGNMENTS

RESULT 1	ADY86938	standard; protein; 1927 AA.
ID	ADY86938	
AC	ADY86938;	
XX		
DT	02-JUN-2005	(first entry)
XX		
DE	Human MRP1-mouse Klr6.2 fusion protein, SEQ ID NO: 1.	
XX		
KM	Ionophore; biosensor; drug screening; diagnostic;	
KW	microorganism detection; potassium channel; fusion protein;	
KX	multidrug resistance protein 1; MRP1; Klr6.2.	
XX		
OS	Homo sapiens.	
OS	Mus musculus.	
OS	Chimeric.	
XX		
FH	Key	Location/Qualifiers
FT	Region	1..1531
FT	Region	/note="Multidrug resistance protein 1 (MRP1)"
FT	Region	1532..1537
FT	Region	/note="Hexaglycine spacer"
FT	Region	1538..1927
FT	Region	/note="Klr6.2 protein"
XX		
XX	US2005063989-A1.	
XX	24-MAR-2005.	
PF	22-SEP-2003; 2003US-00665283.	
XX		
XX	22-SEP-2003; 2003US-00665283.	
XX		
PA	(COMS) COMMISSARIAT ENERGIE ATOMIQUE.	
XX		
XX	Derland R, Garcia E, Prost A, Revilloud J, Vivaudou M;	
XX	WPI; 2005-252611/26.	
XX		
PT	New ion channel hybrid protein, used as electrical sensor for screening	
PT	an agonist/antagonist of a membrane protein and for detecting a	
PT	contaminant/pollutant in a sample.	
XX		
XX	Claim 25; SEQ ID NO 1; 78pp; English.	
CC	The present invention relates to a hybrid protein consisting essentially	

CC of the fusion of a membrane protein with an ion channel which is not
CC naturally coupled to the membrane protein. The hybrid protein is used as
CC an electrical sensor for screening of an agonist/antagonist of a membrane
CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multimeric resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel Kir6.2 protein.

XX Sequence 1927 AA:

Query Match 100.0%; Score 9903; DB 9; Length 1927;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1927; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MALRGFCAGSDSDPLMDNMTMTNPNPTKCFQNTVAVWPCFYLMAKCPFFYLYSRH 60
DB 1 MALRGFCAGSDSDPLMDNMTMTNPNPTKCFQNTVAVWPCFYLMAKCPFFYLYSRH 60
QY 61 DRGYIOMTPLNKTALGFLIMVCMADLFYSFMRSRGIFLAPVFLVSPTLGITLLA 120
DB 61 DRGYIOMTPLNKTALGFLIMVCMADLFYSFMRSRGIFLAPVFLVSPTLGITLLA 120
QY 121 TFLIQLERRKGVQSSGIMLTFLVALVCAALILRSKIMTALKEDAVDLFRDITFYVFS 180
DB 121 TFLIQLERRKGVQSSGIMLTFLVALVCAALILRSKIMTALKEDAVDLFRDITFYVFS 180
QY 181 LLLIOLVSCFSDRSPLFSETHIDPNPCPSSASFLSRITTFWMIIGLIVRGYROPLEGSD 240
DB 181 LLLIOLVSCFSDRSPLFSETHIDPNPCPSSASFLSRITTFWMIIGLIVRGYROPLEGSD 240
QY 241 LMSINKEDTSQOVVPVLVKNMKKECAKTRKQPVKVYVSSKDPAPQKSSKYDANEVEAL 300
DB 241 LMSINKEDTSQOVVPVLVKNMKKECAKTRKQPVKVYVSSKDPAPQKSSKYDANEVEAL 300
QY 301 IVKSPQKWNNSLPKVLKTRGPFLMSFFFKAIHDLMMFSGPQILKLIKRVNDTKAPD 360
DB 301 IVKSPQKWNNSLPKVLKTRGPFLMSFFFKAIHDLMMFSGPQILKLIKRVNDTKAPD 360
QY 361 WQGFYTYLAFYTAACLOTLMVHOYFHCIVSGMRKKTAVIGAVYKALVITNSARKSSTV 420
DB 361 WQGFYTYLAFYTAACLOTLMVHOYFHCIVSGMRKKTAVIGAVYKALVITNSARKSSTV 420
QY 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVILALYLLMLNGPSVLGAVAVMLVPEVN 480
DB 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVILALYLLMLNGPSVLGAVAVMLVPEVN 480
QY 481 AVAMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVILYAMELAFQDKVLAIRQBELKYLK 540
DB 481 AVAMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVILYAMELAFQDKVLAIRQBELKYLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTDENNLILAQTAFFVSLAFNLIIRPLNLP 600
DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTDENNLILAQTAFFVSLAFNLIIRPLNLP 600
QY 601 MVISSIVQASVSLKRLIFLSHEBELPDSIERRPYKDGGSNTSITVRNATFTWASDPT 660
DB 601 MVISSIVQASVSLKRLIFLSHEBELPDSIERRPYKDGGSNTSITVRNATFTWASDPT 660
QY 661 LNGTHFESIPREGALVAVGVQVCGKSSLSALIAEMDKVEGHVAKGSVAVYVQQAIND 720
DB 661 LNGTHFESIPREGALVAVGVQVCGKSSLSALIAEMDKVEGHVAKGSVAVYVQQAIND 720
QY 721 SLRENIILGGCOLLEBYRSVIOACALLPDLLELPSGDREIEGKVNISGGQKQSVLAR 780
DB 721 SLRENIILGGCOLLEBYRSVIOACALLPDLLELPSGDREIEGKVNISGGQKQSVLAR 780
QY 781 AVYSNADIIYLPDDPLSAVDAAHVKHIFENVIGPKMKLNKTRILVTHSKSYLPQVDVIV 840
DB 781 AVYSNADIIYLPDDPLSAVDAAHVKHIFENVIGPKMKLNKTRILVTHSKSYLPQVDVIV 840
QY 841 MSGKISMGSTYOELLARDGAFELRTYASTEOQDAEENGVTGVSQKGAQOMENGM 900
DB 841 MSGKISMGSTYOELLARDGAFELRTYASTEOQDAEENGVTGVSQKGAQOMENGM 900
```

```
DB 841 MSGKISMGSTYOELLARDGAFELRTYASTEOQDAEENGVTGVSQKGAQOMENGM 900
QY 901 LVYDSAGKOLQROUSSSSSYSGDISRHNNSTAELOKAKEETWKLMEADKAQTQYKL 960
DB 901 LVYDSAGKOLQROUSSSSSYSGDISRHNNSTAELOKAKEETWKLMEADKAQTQYKL 960
QY 961 SVYWDYKAIQGLFISFLISIFLPCNNHVSALASNYWLSLWDDPIVNGTOEHTKRLSYG 1020
DB 961 SVYWDYKAIQGLFISFLISIFLPCNNHVSALASNYWLSLWDDPIVNGTOEHTKRLSYG 1020
QY 1021 ALGISOGIAVGYMAVSIIGIILASRCILHVDLHSIILRSPMFPERTPSGNLVNRFSEL 1080
DB 1021 ALGISOGIAVGYMAVSIIGIILASRCILHVDLHSIILRSPMFPERTPSGNLVNRFSEL 1080
QY 1081 DTVDSMTPEVIMKMGSLFNVTGACIYILLATPIAIIIPPLGIYFVGQFVYASSRQL 1140
DB 1081 DTVDSMTPEVIMKMGSLFNVTGACIYILLATPIAIIIPPLGIYFVGQFVYASSRQL 1140
QY 1141 KRLSVSRSPYSHFNETLLGVSVYIRAFEEQERTIHOSDLKVDENOKAYYPSIYANRYLA 1200
DB 1141 KRLSVSRSPYSHFNETLLGVSVYIRAFEEQERTIHOSDLKVDENOKAYYPSIYANRYLA 1200
QY 1201 VRLCVCNGCIYLFALFAVLSRHSLSAGLVGLSVSYSLQVTTYLNMVLVRMSSEMETNIVA 1260
DB 1201 VRLCVCNGCIYLFALFAVLSRHSLSAGLVGLSVSYSLQVTTYLNMVLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFRNYCYLRBDLPVLHINVTINGG 1320
DB 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFRNYCYLRBDLPVLHINVTINGG 1320
QY 1321 EKVGI VERTGAKSSLTGLFRINESABGEIIDGINIAKGLHDLRKTIITIQDPVLF 1380
DB 1321 EKVGI VERTGAKSSLTGLFRINESABGEIIDGINIAKGLHDLRKTIITIQDPVLF 1380
QY 1381 SSGSLRMNLDPSQYSDSEWVTSLELAHKDVSALPDLDHCBAGGENLSVGOROLVCL 1440
DB 1381 SSGSLRMNLDPSQYSDSEWVTSLELAHKDVSALPDLDHCBAGGENLSVGOROLVCL 1440
QY 1441 ARALLRKTKILVLEATPAVDELTDLDIOSTIRQFQEDCTVLTIAHRLNTIMDYTRVIL 1500
DB 1441 ARALLRKTKILVLEATPAVDELTDLDIOSTIRQFQEDCTVLTIAHRLNTIMDYTRVIL 1500
QY 1501 DKGEIOEYGAQSDLLQORGLFYSAKQAGLVGGGGGMLSRKGIIPERYVLTFLAEDPAE 1560
DB 1501 DKGEIOEYGAQSDLLQORGLFYSAKQAGLVGGGGGMLSRKGIIPERYVLTFLAEDPAE 1560
QY 1561 PRYTRERBARFVSKKGCNVAHKNIREQSFLODVFPTLVLDLWMPHTLLIFMSFLCSW 1620
DB 1561 PRYTRERBARFVSKKGCNVAHKNIREQSFLODVFPTLVLDLWMPHTLLIFMSFLCSW 1620
QY 1621 LLFAMVWMLIAFAHGDLPAGEGTNVPCTVSIHSESSAFLFSIYQVTTIGFGRMVTECP 1680
DB 1621 LLFAMVWMLIAFAHGDLPAGEGTNVPCTVSIHSESSAFLFSIYQVTTIGFGRMVTECP 1680
QY 1681 LAIILIVQNIQVGMINAIMGCIEMKTAQAHRAEETLIFSKHAVITLRHRLCFMLRVG 1740
DB 1681 LAIILIVQNIQVGMINAIMGCIEMKTAQAHRAEETLIFSKHAVITLRHRLCFMLRVG 1740
QY 1741 DLKRSMTISATTHQVARKTSPGEVWPLHOUNIPMENGUGNGCILEVAELIYHVHDS 1800
DB 1741 DLKRSMTISATTHQVARKTSPGEVWPLHOUNIPMENGUGNGCILEVAELIYHVHDS 1800
QY 1801 NSPLVDLAPSDLHHHODLEIIVIEGVETGTTIQTARTSYLADEILMGQRFVPIVAED 1860
DB 1801 NSPLVDLAPSDLHHHODLEIIVIEGVETGTTIQTARTSYLADEILMGQRFVPIVAED 1860
QY 1861 GRYSVDYSKFGNTIKVPTPLCTARQDLBDRSLDALTLASSRGLRKRSVAVAKKPKFS 1920
DB 1861 GRYSVDYSKFGNTIKVPTPLCTARQDLBDRSLDALTLASSRGLRKRSVAVAKKPKFS 1920
QY 1921 ISPDLS 1927
DB 1921 ISPDLS 1927
```

RESULT 2	
ADY86943	
ID	ADY86943 standard; protein; 1927 AA.
XX	
AC	ADY86943;
XX	
DT	02-JUN-2005 (first entry)
XX	
DE	MRP1-mouse Kir6.2 mutant (KR370AA) fusion protein, SEQ ID NO: 6.
XX	
KW	Ionophore; biosensor; drug screening; diagnostic;
KM	microorganism detection; potassium channel; fusion protein;
KX	multidrug resistance protein 1; MRP1; Kir6.2; muten.
XX	
OS	Homo sapiens.
OS	Mus musculus.
OS	Chimeric.
XX	
FH	Key
FT	Location/Qualifiers
FT	Region
FT	/note= "Multidrug resistance protein 1 (MRP1)"
FT	1..1531
FT	Region
FT	1532..1537
FT	/note= "Hexaglycine spacer"
FT	1538..1927
FT	Region
FT	/note= "Kir6.2 mutant protein"
FT	1907
FT	Misc-difference
FT	/note= "Wild-type Lys substituted by Ala"
FT	1908
FT	/note= "Wild-type Arg substituted by Ala"
XX	
PN	US2005063989-A1.
PD	
PD	24-MAR-2005.
PE	
PE	22-SEP-2003; 2003US-00665283.
PR	
PR	22-SEP-2003; 2003US-00665283.
XX	
PA	(COMS) COMMISSARIAT ENERGIE ATOMIQUE.
XX	
PI	Derand R, Garcia E, Prost A, Revilloud J, Vivaudou M;
PI	WP1; 2005-252611/26.
PT	
PT	New ion channel hybrid protein, used as electrical sensor for screening
PT	an agonist/antagonist of a membrane protein and for detecting a
PT	contaminant/pollutant in a sample.
XX	
PS	Claim 25; SEQ ID NO 6; 78bp; English.
XX	
CC	The present invention relates to a hybrid protein consisting essentially
CC	of the fusion of a membrane protein with an ion channel which is not
CC	naturally coupled to the membrane protein. The hybrid protein is used as
CC	an electrical sensor for screening of an agonist/antagonist of a membrane
CC	protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC	invention is also useful for medical diagnostics and microorganism
CC	detection. The present sequence is a fusion protein comprising human
CC	channel Kir6.2 mutant (KR370AA) protein.
CC	
CC	Channel Kir6.2 mutant (KR370AA) protein.
XX	
SQ	Sequence 1927 AA:
Query Match	99.9%; Score 9891; DB 9; Length 1927;
Best Local Similarity	99.9%; Pred. No. 0;
Matches 1925; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
OY	1 MALRFGFCADSDPLMDNMVNTWNTSNPDPTKCFONTLVAVWCFFYLMAACEPFYFLYSRH 60
Ddb	1 MALRFGFCADSDPLMDNMVNTWNTSNPDPTKCFONTLVAVWCFFYLMAACEPFYFLYSRH 60
OY	61 DRGVITQMTPLNKTALGFLMLIVCAADLEFVSFWERSRGIFLAPVFLVSPILLGITTLTA 120

Db	1141	KRLBSVSRSPVYSHPNFTLLGVSIVARPEQERFIHQSDLKVYDENQKAYPSPSIANRWLA	1200
Qy	1201	VRLECVGNCTIVLPAALFAVIVSRHSISAGLVGLSVSYSLQVTTYLMNLVRMSSEMETNIVA	1260
Db	1201	VRLECVGNCTIVLPAALFAVIVSRHSISAGLVGLSVSYSLQVTTYLMNLVRMSSEMETNIVA	1260
Qy	1261	VERLKEYSETEKEAPMQLQETAPPSWQVGRVERFRNCLTRYREDLDVLIHINVTNGG	1320
Db	1261	VERLKEYSETEKEAPMQLQETAPPSWQVGRVERFRNCLTRYREDLDVLIHINVTNGG	1320
Qy	1321	EKVGIVGTGAGKSSLTGLGRFINESSAGEEIIIDGINIAKIGLHDLRFKTIITIPQDPVLF	1380
Db	1321	EKVGIVGTGAGKSSLTGLGRFINESSAGEEIIIDGINIAKIGLHDLRFKTIITIPQDPVLF	1380
Qy	1381	SGSLRPMNIDPPSOYSDEEVTWTSLELAHKDPVPSALPDLKDHCECAGGNNLSVGRQIVCL	1440
Db	1381	SGSLRPMNIDPPSOYSDEEVTWTSLELAHKDPVPSALPDLKDHCECAGGNNLSVGRQIVCL	1440
Qy	1441	ARALLRKTKIIVLDBATAVADLETBDLIQSTRTQFEDCTVLTIIHRIANTIMDYTRVVL	1500
Db	1441	ARALLRKTKIIVLDBATAVADLETBDLIQSTRTQFEDCTVLTIIHRIANTIMDYTRVVL	1500
Qy	1501	DKGEIOEYGAESDILLQORGLFYSMAKDAAGLVGGGGGMLSRKGIIPEEYVITRLAEDPAE	1560
Db	1501	DKGEIOEYGAESDILLQORGLFYSMAKDAAGLVGGGGGMLSRKGIIPEEYVITRLAEDPAE	1560
Qy	1561	PRYRFRERARAFVSKKGNCAVNAHKRIREQGRPLDOVFTTVLDKMPHLLIFPMSCFSW	1620
Db	1561	PRYRFRERARAFVSKKGNCAVNAHKRIREQGRPLDOVFTTVLDKMPHLLIFPMSCFSW	1620
Qy	1621	LLFANVWMLIAFAHGDLPAGEGTNPVCYTSIHSFSSAEFLSIEVOTYIGFGGRMWTBEP	1680
Db	1621	LLFANVWMLIAFAHGDLPAGEGTNPVCYTSIHSFSSAEFLSIEVOTYIGFGGRMWTBEP	1680
Qy	1681	LAILLIVQNIIVGLMINAIMLGCIIPMKTQAQAHRAETLLIFSKHAVITLLRHGSLCEMLFVG	1740
Db	1681	LAILLIVQNIIVGLMINAIMLGCIIPMKTQAQAHRAETLLIFSKHAVITLLRHGSLCEMLFVG	1740
Qy	1741	DLRKSMTISATIHMQVARKTSPBEEVVP/LHNOVDPMENGVGAGNCEIPLVAPLIITHVIDS	1800
Db	1741	DLRKSMTISATIHMQVARKTSPBEEVVP/LHNOVDPMENGVGAGNCEIPLVAPLIITHVIDS	1800
Qy	1801	NSPLVDLAPSDLHHHQDEIIVILEGVEVETGITTQARTSYLADEIILMGORFVPIVAEED	1860
Db	1801	NSPLVDLAPSDLHHHQDEIIVILEGVEVETGITTQARTSYLADEIILMGORFVPIVAEED	1860
Qy	1861	GRYSVDYSKFGNTITKVPTPLCTARQJLDBDRSLLDALTLASSRGLPRKDSVAVAKAKPKFS	1920
Db	1861	GRYSVDYSKFGNTITKVPTPLCTARQJLDBDRSLLDALTLASSRGLPRKDSVAVAKAKPKFS	1920
Qy	1921	ISPDLSL 1927	
Db	1921	ISPDLSL 1927	

ADY86945	ADY86945 standard; protein; 1947 AA.
ID	
XX	
AC	ADY86945;
XX	
DT	02-JUN-2005 (first entry)
XX	
DE	Human MRP1-mouse Kir6.2-HA fusion protein, SEQ ID NO: 8.
XX	
KW	Ionophore; biosensor; drug screening; diagnostic;
KW	microorganism detection; potassium channel; fusion protein;
KW	multidrug resistance protein 1; MRP1; Kir6.2.
XX	
OS	Homo sapiens.
OS	Mus musculus.
OS	Chimeric.
OS	Unidentified.

XX	Key	Location/Qualifiers
FH	Region	1..1531
FT	Region	/note= "Multidrug resistance protein 1 (MRP1)"
FT	Region	1532..1537
FT	Region	/note= "Hexaglycine spacer"
FT	Region	1538..1947
FT	Region	/note= "Kir6.2 protein containing HA epitope peptide"
FT	Region	1651..1659
FT	Region	/note= "HA epitope peptide"
XX		
PN		US2005063989-A1.
XX		
PD		24-MAR-2005.
XX		
PX		22-SEP-2003; 2003US-00665283.
PR		22-SEP-2003; 2003US-00665283.
XX		
PA	(COMS) COMMISSARIAT ENERGIE ATOMIQUE.	
XX	Derland R, Garcia E, Prost A, Revilloud J, Vivaudou M;	
PI	WPI; 2005-252611/26.	
DR		
XX	New ion channel hybrid protein, used as electrical sensor for screening	
PT	an agonist/antagonist of a membrane protein and for detecting a	
XX	contaminant/pollutant in a sample.	
PS		
XX	Claim 25; SEQ ID NO 8; 78pp; English.	
CC	The present invention relates to a hybrid protein consisting essentially	
CC	of the fusion of a membrane protein with an ion channel which is not	
CC	naturally coupled to the membrane protein. The hybrid protein is used as	
CC	an electrical sensor for screening of an agonist/antagonist of a membrane	
CC	protein, drugs and for detecting a contaminant/pollutant in a sample. The	
CC	invention is also useful for medical diagnostics and microorganism	
CC	detection. The present sequence is a fusion protein comprising human	
CC	multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium	
CC	channel Kir6.2 protein containing HA epitope peptide.	
SQ	Sequence 1947 AA:	
	Query Match	99.7%; Score 9873; DB 9; Length 1947;
	Best Local Similarity	99.0%; Pred. No. 0;
	Matches 1927; Conservative	0; Mismatches 0; Indels 20; Gaps 2
OY	1 MALRFGCADSDPLMDNMVMTNMTSNPFTKCFONTLVVMWPCEFLMACFPFFLYLSRH	60
DB		
	1 MALRFECGADSDPLMDNMVMTNMTSNPFTKCFONTLVVMWPCEFLMACFPFFLYLSRH	60
OY	61 DRGYIQMPLNKTKTALGFLLIMIVCMADLFYSFWERSKGIFLAPFVLSPTLTGITTTLA	120
DB	61 DRGYIQMPLNKTKTALGFLLIMIVCMADLFYSFWERSKGIFLAPFVLSPTLTGITTTLA	120
OY	121 TFLIOLERRKGVSSGIMLTFWLVALVCALAIIRSKIMTALKEDAQVDLFRDITTYVFS	180
DB	121 TFLIOLERRKGVSSGIMLTFWLVALVCALAIIRSKIMTALKEDAQVDLFRDITTYVFS	180
OY	181 LLLIOLVISCDFSRPLFSETHIDNPCESSASFLSIITFMWIGLYRGROPLEGSD	240
DB	181 LLLIOLVISCDFSRPLFSETHIDNPCESSASFLSIITFMWIGLYRGROPLEGSD	240
OY	241 LMSLNKEDETSBOVVVLVKNMKKECAKTRKQPVKVYVSCKDAQCRESSKRVANEEVEAL	300
DB	241 LMSLNKEDETSBOVVVLVKNMKKECAKTRKQPVKVYVSCKDAQCRESSKRVANEEVEL	300
OY	301 IYKSPOKEWNPSLFFVLYKTGGPYFLMSGFFPKAIDLMWFSGPOLIKLLIKFVNDTKAD	360
DB	301 IYKSPOKEWNPSLFFVLYKTGGPYFLMSGFFPKAIDLMWFSGPOLIKLLIKFVNDTKAD	360
OY	361 WGGFYTVALLPFTACLOTLVLHQYRHICVGSGRKRTAVIGAVYRKALVTNTSAKRSSTV	420
DB	361 WGGFYTVALLPFTACLOTLVLHQYRHICVGSGRKRTAVIGAVYRKALVTNTSAKRSSTV	420

[illegible]

Qy	1501	DKGEIIOEYCAPSDLLQORLFLFSMADAGLVGGGGGGM.SRKIIIPBEYVLTALAEPAE	1560
Db	1501	DKGEIOEYCAPSDLLQORLFLFSMADAGLVGGGGGGM.SRKGIIPBEYVLTALAEPAE	1560
Qy	1561	PYRFRERARFVSKKGCNVAAHKNIREQGRFLQDVFTTLVLDKMPHTLLIFTMSFLCSW	1620
Db	1561	PYRFRERARFVSKKGCNVAAHKNIREQGRFLQDVFTTLVLDKMPHTLLIFTMSFLCSW	1620
Qy	1621	ILFAMVWMLIAEPAHG-----DLAP-----GEGTNVPCVTSIHSSSAFLF	1660
Db	1621	ILFAMVWMLIAEPAHGDLVAYMEKGIITDLAPYPYDVPDYAGEGTNVPCVTSIHSSSAFLF	1660
Qy	1661	SIEVQVTITGFGGRMTTEECPLAILLIVONIYGLMNAIMLGCIIPKTAQAHRAETLIF	1720
Db	1681	SIEVQVTITGFGGRMTTEECPLAILLIVONIYGLMNAIMLGCIIPKTAQAHRAETLIF	1740
Qy	1721	SGHAAYTLRHGRLCFMLRVGDLRKSMIISATIMQVVRKTTSPERGEVPLHQVDIPWENG	1780
Db	1741	SGHAAYTLRHGRLCFMLRVGDLRKSMIISATIMQVVRKTTSPERGEVPLHQVDIPWENG	1800
Qy	1781	VGGNGIFLVAAPLIIVHVIDSNSPLYDLAPSDLHHQDLBIIVLGEVETTGITTOARTS	1840
Db	1801	VGGNGIFLVAAPLIIVHVIDSNSPLYDLAPSDLHHQDLBIIVLGEVETTGITTOARTS	1860
Qy	1841	YLADIELLQGRVPIVAEEDGRYSVDYSKFGNTIKVPTPLCTAROLDDEBSLIDLALTLAS	1900
Db	1861	YLADIELLQGRVPIVAEEDGRYSVDYSKFGNTIKVPTPLCTAROLDDEBSLIDLALTLAS	1920
Qy	1901	SRGPLRKRSVAAYAKAKPKFSISPDLSL	1927
Db	1921	SRGPLRKRSVAAYAKAKPKFSISPDLSL	1947
RESULT 4			
ADY86941			
ID	ADY86941	standard; protein; 1891 AA.	
XX	ADY86941;		
AC			
XX			
DT	02-JUN-2005	(first entry)	
XX			
DE	Human MRP1-Kir6.2 mutant delctac36 fusion protein, SEQ ID NO: 4.		
XX			
KM	Ionophore, biosensor; drug screening; diagnostic;		
KW	microorganism detection; potassium channel; fusion protein;		
KM	multidrug resistance protein 1; MRP1; Kir6.2; mutcin.		
XX			
OS	Homo sapiens.		
OS	Mus musculus.		
OS	Chimeric.		
XX			
FH	Key	Location/Qualifiers	
FT	Region	1..1531	
FT	Region	/note= "Multidrug resistance protein 1 (MRP1) "	
FT	Region	1532..1537	
FT	Region	/note= "Hexaglycine spacer"	
FT	Region	1538..1891	
FT	Region	/note= "Kir6.2 mutant delctac36 protein"	
XX			
PN	US2005063989-A1.		
XX			
PD	24-MAR-2005.		
XX			
PF	22-SEP-2003; 2003US-00665283.		
XX			
PR	22-SEP-2003; 2003US-00665283.		
XX			
PA	(COMS) COMMISSARIAT ENERGIE ATOMIQUE.		
XX			
PI	Derand R, Garcia E, Prost A, Revilloud J, Vivaudou M;		
XX			
DR	WPI; 2005-252611/26.		

XX New ion channel hybrid protein, used as electrical sensor for screening
PT an agonist/antagonist of a membrane protein and for detecting a
PT contaminant/pollutant in a sample.
XX
PS Claim 25; SEQ ID NO 4; 78pp; English.
CC The present invention relates to a hybrid protein consisting essentially
CC of the fusion of a membrane protein with an ion channel which is not
CC naturally coupled to the membrane protein. The hybrid protein is used as
CC an electrical sensor for screening of an agonist/antagonist of a membrane
CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel Kir6.2 mutant detac36 protein.
XX
SQ Sequence 1891 AA;
Query Match 98.3%; Score 9734; DB 9; Length 1891;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALGFCSDGSDPLMDNMNTMTNSNPDFTKCFQNTVLWVPCFYLMACFPFYLYSRH 60
DB 1 MALGFCSDGSDPLMDNMNTMTNSNPDFTKCFQNTVLWVPCFYLMACFPFYLYSRH 60
QY 61 DRGYIOMTPLNKRTALGFLIMIVCMADLFYSFWERSNGIFLAPYFLVSPTLGITTTLLA 120
DB 61 DRGYIOMTPLNKRTALGFLIMIVCMADLFYSFWERSNGIFLAPYFLVSPTLGITTTLLA 120
QY 121 TFLIQLEERRKGVSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVFS 180
DB 121 TFLIQLEERRKGVSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVFS 180
QY 181 LLLIQVLVSCSDSRLFSSTIHDNPPCESASLSLITFWMTGLVGRGROLESBD 240
DB 181 LLLIQVLVSCSDSRLFSSTIHDNPPCESASLSLITFWMTGLVGRGROLESBD 240
QY 181 LLLIQVLVSCSDSRLFSSTIHDNPPCESASLSLITFWMTGLVGRGROLESBD 240
DB 181 LLLIQVLVSCSDSRLFSSTIHDNPPCESASLSLITFWMTGLVGRGROLESBD 240
QY 241 LMSLNKEDTSQOVVPLVKNMKKECAKTRKOPVKVYSSKDPACPKESSKYDANEVEAL 300
DB 241 LMSLNKEDTSQOVVPLVKNMKKECAKTRKOPVKVYSSKDPACPKESSKYDANEVEAL 300
QY 301 IVKSPQKEMNPSLFKVLVYKTFGPYPLMSFPFKAHDLMMFSGPOLIKLIFVNDTKAPD 360
DB 301 IVKSPQKEMNPSLFKVLVYKTFGPYPLMSFPFKAHDLMMFSGPOLIKLIFVNDTKAPD 360
QY 301 IVKSPQKEMNPSLFKVLVYKTFGPYPLMSFPFKAHDLMMFSGPOLIKLIFVNDTKAPD 360
DB 301 IVKSPQKEMNPSLFKVLVYKTFGPYPLMSFPFKAHDLMMFSGPOLIKLIFVNDTKAPD 360
QY 361 WQGYFYTLVFTVACLOTLVLAHQYHICFVSGMRKKTAVIGAVYKKAIVITNSAKSSSTV 420
DB 361 WQGYFYTLVFTVACLOTLVLAHQYHICFVSGMRKKTAVIGAVYKKAIVITNSAKSSSTV 420
QY 361 WQGYFYTLVFTVACLOTLVLAHQYHICFVSGMRKKTAVIGAVYKKAIVITNSAKSSSTV 420
DB 361 WQGYFYTLVFTVACLOTLVLAHQYHICFVSGMRKKTAVIGAVYKKAIVITNSAKSSSTV 420
QY 421 GEIYNLMSVDKORFMDLATYINMISAPLOVTLALYLMNLGSPVLAVGAVMVLMPVN 480
DB 421 GEIYNLMSVDKORFMDLATYINMISAPLOVTLALYLMNLGSPVLAVGAVMVLMPVN 480
QY 481 AVMAKTKTYOVAAHMKSKDNRIKLMNEILNGIKVLKLYAMELAFKQVLAIRQEBLKVLK 540
DB 481 AVMAKTKTYOVAAHMKSKDNRIKLMNEILNGIKVLKLYAMELAFKQVLAIRQEBLKVLK 540
QY 541 KSAIYLSAVGTPTWCTPFLVALCTPAVYVTTDENNILLDAQTAFFVSIALFNILRPLNTLP 600
DB 541 KSAIYLSAVGTPTWCTPFLVALCTPAVYVTTDENNILLDAQTAFFVSIALFNILRPLNTLP 600
QY 601 MVISIYVASVSLKRLIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
DB 601 MVISIYVASVSLKRLIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
QY 661 LINGITFSIPBEGALVAVGVGGKSSLSALLAEMDKVEGHVAIKGSVAVYVQQAQMIQND 720
DB 661 LINGITFSIPBEGALVAVGVGGKSSLSALLAEMDKVEGHVAIKGSVAVYVQQAQMIQND 720
QY 721 SLRENILFGQLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKVNISSGQKORVSLAR 780
DB 721 SLRENILFGQLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKVNISSGQKORVSLAR 780

DB 721 SLRENILFGQLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKVNISSGQKORVSLAR 780
QY 781 AVYSNADIYLEDPLSAVDAAVGHKIFENVIGPKMLKNKTRIIIVTHSMSTYLPQVDVITV 840
DB 781 AVYSNADIYLEDPLSAVDAAVGHKIFENVIGPKMLKNKTRIIIVTHSMSTYLPQVDVITV 840
QY 841 MSGKISBMSYQGLLARDGAFAFLRTYASTBEOBDAEENGVTGVSQPGKEAQMENG 900
DB 841 MSGKISBMSYQGLLARDGAFAFLRTYASTBEOBDAEENGVTGVSQPGKEAQMENG 900
QY 901 LVTSAGKQLOKROUSSSSSYSGDISRHHNSTAELOKAEKEEFTWKLMEADKAQGVYKL 960
DB 901 LVTSAGKQLOKROUSSSSSYSGDISRHHNSTAELOKAEKEEFTWKLMEADKAQGVYKL 960
QY 961 SVYWDYKAIGLFSIFLFLMCHNVSAIASNYMSLMTDPIVNGTQEHKTVRLSYVG 1020
DB 961 SVYWDYKAIGLFSIFLFLMCHNVSAIASNYMSLMTDPIVNGTQEHKTVRLSYVG 1020
QY 1021 ALGISQGIYAVRGYMAVSIIGILASRCLHVDLHSLRSPMSFPERTPSGNLVNRFSKEL 1080
DB 1021 ALGISQGIYAVRGYMAVSIIGILASRCLHVDLHSLRSPMSFPERTPSGNLVNRFSKEL 1080
QY 1081 DTVDSMTPEVTKMFMGSLFENVIGACIVILLATPIAIIIPPLGLIYFVQGFVYASSRQL 1140
DB 1081 DTVDSMTPEVTKMFMGSLFENVIGACIVILLATPIAIIIPPLGLIYFVQGFVYASSRQL 1140
QY 1141 KRLSVSRSPVYSHFNETLLGVSVIRAEQGERFIHOSDLKVDENOKAYYPSIVANRWLA 1200
DB 1141 KRLSVSRSPVYSHFNETLLGVSVIRAEQGERFIHOSDLKVDENOKAYYPSIVANRWLA 1200
QY 1201 VRLBCVGNCTYLPAALPAVIRSHLSAGLVCLSVYSISQVTTYINMLVYRSMSEMETNIVA 1260
DB 1201 VRLBCVGNCTYLPAALPAVIRSHLSAGLVCLSVYSISQVTTYINMLVYRSMSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIQTETAPPSMPOQVRAVEFRNYCLRYREDLDPVLHINVTINGG 1320
DB 1261 VERLKEYSETEKEAPWQIQTETAPPSMPOQVRAVEFRNYCLRYREDLDPVLHINVTINGG 1320
QY 1321 EKVGIYGRGTGAKSSLTLGLFRINESAGEIIIDGINIAKIGLHDLRFXITIIIPQDPVLF 1380
DB 1321 EKVGIYGRGTGAKSSLTLGLFRINESAGEIIIDGINIAKIGLHDLRFXITIIIPQDPVLF 1380
QY 1381 SGSLRMLNDPFSQYSDSEEWMTSLBLAHKDPVSALPDLDHCEKAGEGNSLVGQROLVCL 1440
DB 1381 SGSLRMLNDPFSQYSDSEEWMTSLBLAHKDPVSALPDLDHCEKAGEGNSLVGQROLVCL 1440
QY 1441 ARLARKTKIIVLEBATAVDLETDLQSTIFRQFEDCTVLTAAHRLNTIMDYTRVYL 1500
DB 1441 ARLARKTKIIVLEBATAVDLETDLQSTIFRQFEDCTVLTAAHRLNTIMDYTRVYL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYMAKQAGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAE 1560
DB 1501 DKGEIOEYGAPSDLLQORGLFYMAKQAGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAE 1560
QY 1561 PRYTRERRARFVSKGNCNVAHKNIREQGRFLQDVFTTLVDLKKPHULLFTMSFLCSW 1620
DB 1561 PRYTRERRARFVSKGNCNVAHKNIREQGRFLQDVFTTLVDLKKPHULLFTMSFLCSW 1620
QY 1621 LFPFVWMLIAFAHGDLAPGEGTNVPCVTSIHSFSSALFESI EVOVTTIGFGGRMTEBCP 1680
DB 1621 LFPFVWMLIAFAHGDLAPGEGTNVPCVTSIHSFSSALFESI EVOVTTIGFGGRMTEBCP 1680
QY 1681 LAIILLYQNTVGLMINAIMGCIIFMKTQAHRRAETLIFSKHAVITLRRHGLCFMLERVG 1740
DB 1681 LAIILLYQNTVGLMINAIMGCIIFMKTQAHRRAETLIFSKHAVITLRRHGLCFMLERVG 1740
QY 1741 DLKSMITISATIHQVAKTSSPEGEVPLHQVDIPMENGVGANGIPIVAPLIIYHVIDS 1800
DB 1741 DLKSMITISATIHQVAKTSSPEGEVPLHQVDIPMENGVGANGIPIVAPLIIYHVIDS 1800
QY 1801 NSPLVDLAPSDLHHQDLEIIVILEGVETTGITTOATSYLADEIILMGORFVPIVAED 1860
DB 1801 NSPLVDLAPSDLHHQDLEIIVILEGVETTGITTOATSYLADEIILMGORFVPIVAED 1860

QY 1861 GRYSVDYSKFGNTIKVPTPLCTARQLDDEDRS 1891
 DB 1861 GRYSVDYSKFGNTIKVPTPLCTARQLDDEDRS 1891

RESULT 5
 AAM57486
 ID AAM57486 standard; protein, 1531 AA.
 AC AAM57486;
 XX
 XX 14-AUG-1998 (first entry)
 XX
 XX Human MRP variant 1cPgpa (Lei/Pgpa).
 XX
 XX Multidrug resistance-associated protein; MRP; tumour; human; variant;
 KM multidrug resistance; MDR; leishmania P-glycoprotein; 1cPgpa; Lei/Pgpa.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 685
 FT /label= L685S
 FT /note= "wild-type Leu is replaced by Ser"
 FT Misc-difference 1282
 FT /label= R1282A
 FT /note= "wild-type Arg is replaced by Ala"
 FT
 XX US5766880-A.
 XX
 PD 16-JUN-1998.
 XX
 PF 05-JUN-1995; 95US-00463092.
 XX
 XX 27-OCT-1992; 92US-00966923.
 PR 08-MAR-1993; 93US-00029340.
 PR 26-OCT-1993; 93US-00141893.
 PR 20-MAR-1995; 95US-00407207.
 XX
 PA (TOOH) UNIV QUEBENS KINGSTON.
 XX
 PI Cole SP, Deeley RG;
 XX
 DR WPI: 1998-361687/31.
 DR N-PSDB; AAV31498.
 XX
 PT DNA encoding protein associated with multi-drug resistance - useful for
 PT as probe for identifying multi-drug resistant tumour cells.
 XX
 PS Claim 1; Col 67-78; 82pp; English.
 XX
 CC This represents a variant of the human multidrug resistance-associated
 CC protein (MRP). This natural variant is a leishmania P-glycoprotein related
 CC molecule 1cPgpa (Lei/Pgpa). The human and murine MRP nucleic acid
 CC molecules can be used as probes for identifying multidrug resistant
 CC tumour cells by hybridisation to mRNA from tumour cells. The antisense
 CC nucleic acid can be used to reverse multidrug resistance (MDR). A
 CC recombinant expression vector containing the MRP nucleic acid molecules
 CC operatively linked to at least one regulatory sequence can be used to
 CC transform a host cell to produce a recombinant MDR-associated protein
 XX
 SQ Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 2; Length 1531;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDLLMNMNTMNTSNPDFTGCGNTYLVWPCRYLMACPFYFLYLSRH 60
 DB 1 MALRGFCSADGSDLLMNMNTMNTSNPDFTGCGNTYLVWPCRYLMACPFYFLYLSRH 60

QY 61 DRGYIQMTPLNKTKTALGFLIMIVCMADLFYSFWEBSRGIFLAVFLVSPFLIGITTLA 120

DB 61 DRGYIQMTPLNKTKTALGFLIMIVCMADLFYSFWEBSRGIFLAVFLVSPFLIGITTLA 120
 QY 121 TFLIQLERKKVQSSGIMLTFWVALYCALALILSKIMTALKEDAOVDLFFDITFYVFS 180
 DB 121 TFLIQLERKKVQSSGIMLTFWVALYCALALILSKIMTALKEDAOVDLFFDITFYVFS 180
 QY 181 LLLIQLVLSGSDSPLESETIHDNPNCPSSSASFRLRTFWMTGLIVRGYRQPLESSD 240
 DB 181 LLLIQLVLSGSDSPLESETIHDNPNCPSSSASFRLRTFWMTGLIVRGYRQPLESSD 240
 QY 241 LMSLNKEDTSEQVPLVVKWKKCECAKTRKQPVVYVSSKDPAPKSSSKVDANEVEAL 300
 DB 241 LMSLNKEDTSEQVPLVVKWKKCECAKTRKQPVVYVSSKDPAPKSSSKVDANEVEAL 300
 QY 301 IVKSPQKEMNPSLFKVLKFTGPFYFLMSFFFKAIHDLMSFQIILKILKFVNDTKAD 360
 DB 301 IVKSPQKEMNPSLFKVLKFTGPFYFLMSFFFKAIHDLMSFQIILKILKFVNDTKAD 360
 QY 361 WQGYFYTVLLFVTAQLOTVLVHOYFHI CFSVGMRIKTAIVGAVRKALVITNSARKSTV 420
 DB 361 WQGYFYTVLLFVTAQLOTVLVHOYFHI CFSVGMRIKTAIVGAVRKALVITNSARKSTV 420
 QY 421 GEIVNLMSVDAQRFMDLATYINMWSAPLOVITLALYLLMNLGSPVLAGVAVMTAMYPVN 480
 DB 421 GEIVNLMSVDAQRFMDLATYINMWSAPLOVITLALYLLMNLGSPVLAGVAVMTAMYPVN 480
 QY 481 AVMAKTKTYQVAMKSKDNRIKLMELINGIKVLKYAMELAKDYALIROBELKVLK 540
 DB 481 AVMAKTKTYQVAMKSKDNRIKLMELINGIKVLKYAMELAKDYALIROBELKVLK 540
 QY 541 KSAVLASVGTFTWCTPFVALCTFPAVYVITDENNII DAOQTAFSIALFNILRPPLILP 600
 DB 541 KSAVLASVGTFTWCTPFVALCTFPAVYVITDENNII DAOQTAFSIALFNILRPPLILP 600
 QY 601 MVISSIVQASVSLKRLIFLSHEELBPSIERRPVKDGGGNSITVNAATFTMARSDPT 660
 DB 601 MVISSIVQASVSLKRLIFLSHEELBPSIERRPVKDGGGNSITVNAATFTMARSDPT 660
 QY 661 LINGITFSIPGALVAVGVGCGKSSLSALAEMLDMRVBHVAIKGSVAVVPQAMQND 720
 DB 661 LINGITFSIPGALVAVGVGCGKSSLSALAEMLDMRVBHVAIKGSVAVVPQAMQND 720
 QY 721 SLRENILFGCOLBEPYRSTYIOACALLPDLILPDSGRTEIGEKVNLSSGQKORVSLAR 780
 DB 721 SLRENILFGCOLBEPYRSTYIOACALLPDLILPDSGRTEIGEKVNLSSGQKORVSLAR 780
 QY 781 AVYSNADIYLFDDPLASVDAHVGHIFENYIGPKGMKNTRILIVTHSMGYLPQVDYIIV 840
 DB 781 AVYSNADIYLFDDPLASVDAHVGHIFENYIGPKGMKNTRILIVTHSMGYLPQVDYIIV 840
 QY 841 MSGGKISEMSYOELARDAFAEFLRTVASTEOQDABENGVGVSGPKGAKOMENGM 900
 DB 841 MSGGKISEMSYOELARDAFAEFLRTVASTEOQDABENGVGVSGPKGAKOMENGM 900
 QY 901 LVYDSAGKQIQRQLSSSSSSYSGDISRRHNSTABLOKAEAKKETWKLMADKQOTGVKL 960
 DB 901 LVYDSAGKQIQRQLSSSSSSYSGDISRRHNSTABLOKAEAKKETWKLMADKQOTGVKL 960
 QY 961 SVYWDYWKALGLFSFLSTFLPMCNHVSALASNYWLSIMTDDPIVNGTOHHTVRLSVYG 1020
 DB 961 SVYWDYWKALGLFSFLSTFLPMCNHVSALASNYWLSIMTDDPIVNGTOHHTVRLSVYG 1020
 QY 1021 ALGISQGIADVGVSMVSIIGIILASRLAHVLDLSILRSPMSFEPSPGNLVNRFSEKL 1080
 DB 1021 ALGISQGIADVGVSMVSIIGIILASRLAHVLDLSILRSPMSFEPSPGNLVNRFSEKL 1080
 QY 1081 DTVDSMTPEYIKMFMGSLFNVIGACIVILATPLAIIIPPLGIYFFVQRFVYASSROL 1140
 DB 1081 DTVDSMTPEYIKMFMGSLFNVIGACIVILATPLAIIIPPLGIYFFVQRFVYASSROL 1140
 QY 1141 KRLSEVSRSPVYSHFNETLLGVSVIRAFEEQSEFFIHOSDLVDENOKAYPSIVANRWLA 1200

Db 1141 KRLSEVSRSPVYSHNETLGVSVIRAEEDRFTIHGSDLKVDENOKAYVPSIVANKRLA 1200
 QY 1201 VRLCEVCNACIVLFAALFAVISRHSLSAGLVGLSVSYLQVTTYTNMLVRMSSEMETNIVA 1260
 Db 1201 VRLCEVCNACIVLFAALFAVISRHSLSAGLVGLSVSYLQVTTYTNMLVRMSSEMETNIVA 1260
 QY 1261 VERLKEYSETKEKAPWQIOETAPPSRSMQVGRVERRANCLARYEDLDVLRHINVTINGG 1320
 Db 1261 VERLKEYSETKEKAPWQIOETAPPSRSMQVGRVERRANCLARYEDLDVLRHINVTINGG 1320
 QY 1321 EKVGIVRGTAGKSLTGLFRINSEAGEIIIDININAKIGLHDLRFKTIIPDPVLF 1380
 Db 1321 EKVGIVRGTAGKSLTGLFRINSEAGEIIIDININAKIGLHDLRFKTIIPDPVLF 1380
 QY 1381 SGLSRMNLDPFSPQYSDSEEWTSLELAHLKDFVSALPDKLDHBCAGGEMLSVGQRLVCL 1440
 Db 1381 SGLSRMNLDPFSPQYSDSEEWTSLELAHLKDFVSALPDKLDHBCAGGEMLSVGQRLVCL 1440
 QY 1441 ARALLRKTIIVDEBATAVVDLETDLLIQSTIRTPQFECTVLTNAHRLNTIMDTIRVVL 1500
 Db 1441 ARALLRKTIIVDEBATAVVDLETDLLIQSTIRTPQFECTVLTNAHRLNTIMDTIRVVL 1500
 QY 1501 DKGEIOEXGAPSDLLQQRGLFYSMKADGLV 1531
 Db 1501 DKGEIOEXGAPSDLLQQRGLFYSMKADGLV 1531

RESULT 6

AAM74471
 ID AAM74471 standard; protein; 1531 AA.

AC AAM74471;

DT 18-MAY-1999 (first entry)

DE Human multidrug resistance-associated protein variant.

KM Multidrug resistance-associated protein; MDR; human; diagnosis;

OS Homo sapiens.

XX US5882875-A.

PD 16-MAR-1999.

PF 05-JUN-1995; 95US-00462109.

FR 27-OCT-1992; 92US-00966923.

PR 08-MAR-1993; 93US-00029340.

PR 26-OCT-1993; 93US-00141893.

PR 20-MAR-1995; 95US-00407207.

PA (TOOH) UNIV QUEBENS KINGSTON.

P1 Cole SPC, Deeley RG;

DR WPI; 1999-214061/18.

DR N-PSDB; AAX21977.

PT Identifying a multidrug resistant tumor cell by contacting the cell with

PT an antibody/antigen-binding fragment - which binds to an expressed

PT protein encoded by multidrug resistance-associated protein (MRP) nucleic

PT acid.

PS Claim 3; Col 69-80; 80pp; English.

CC This sequence is the human multidrug resistance-associated (MDR) protein.

CC The invention relates to a method for identifying a multidrug resistant

CC (MDR) tumor cell. Compositions and methods utilizing the MDR proteins

CC can be used to treat patients with tumors displaying multidrug

CC resistance, particularly those displaying resistance to antineoplastic

CC epipodophyllotoxins, vinca alkaloids, and hydrophobic drugs. The methods

CC for inhibiting/killing a MDR tumor cell can be useful for treating
 CC breast cancer, leukemias, fibrosarcomas, cervical cancer, gliomas,
 CC thymomas, neuroblastomas and lung cancer. The MDR DNA sequences when
 CC labeled are useful as molecular probes for diagnosing multidrug
 CC resistance of a tumor (using cells from a tumour biopsy) and for
 CC designing ribozymes which are capable of cleaving a single-stranded
 CC nucleic acid encoding a protein having MRP activity. Recombinant
 CC expression vectors containing human MDR coding sequences can be
 CC transfected into a drug sensitive cell line to produce a protein in the
 CC cell which confers MDR, protecting non-resistant non-tumour cells from
 CC the effects of chemotherapeutics has major clinical importance. Cells
 CC transformed with the MDR coding sequences are useful for testing
 CC potential therapeutic agents for their effectiveness against MDR cells
 CC and for identifying chemosensitizers of a therapeutic agent

SQ Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 2; Length 1531;

Best local similarity 100.0%; Pred. NO. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1531; Conservative 0;

QY 1 MALRGFCSADGSDPLMDMNVMTNTPDKPQNTVWVPCFYLMACFPFELYLSRH 60

Db 1 MALRGFCSADGSDPLMDMNVMTNTPDKPQNTVWVPCFYLMACFPFELYLSRH 60

QY 61 DRGYIOMTPINKTKTALGFLIMVCMADLFYSFWEBSRGIFLAVFLVSPFLGTTLLA 120

Db 61 DRGYIOMTPINKTKTALGFLIMVCMADLFYSFWEBSRGIFLAVFLVSPFLGTTLLA 120

QY 121 TELIOLERRKGVSSGIMTFEVLVALCALALIRSKMTAKEDAOVDLPFDIFYYVFS 180

Db 121 TELIOLERRKGVSSGIMTFEVLVALCALALIRSKMTAKEDAOVDLPFDIFYYVFS 180

QY 181 LLLIQLVLSGSDSPFLFSETIHDNCPBESSASFLSRITPFWITGLIVRGYROPLESGD 240

Db 181 LLLIQLVLSGSDSPFLFSETIHDNCPBESSASFLSRITPFWITGLIVRGYROPLESGD 240

QY 241 LMSINKEDTSQOVVPIVKNKKCEAKTRKOPVAVYVSSKDPAPKRESSKYDANEVEAL 300

Db 241 LMSINKEDTSQOVVPIVKNKKCEAKTRKOPVAVYVSSKDPAPKRESSKYDANEVEAL 300

QY 301 IVKSPQKEMNPSLTKULYKTFGPFELMSFPFKALHDLMFSGPOLIKLILFVNDTKAPD 360

Db 301 IVKSPQKEMNPSLTKULYKTFGPFELMSFPFKALHDLMFSGPOLIKLILFVNDTKAPD 360

QY 361 WQGYFYTVLFTVTAQCLQTLVHOFYHICFVSGMRKTAIVAGVYRKALVITNSARKSSTV 420

Db 361 WQGYFYTVLFTVTAQCLQTLVHOFYHICFVSGMRKTAIVAGVYRKALVITNSARKSSTV 420

QY 421 GEIVNLSVDAQRPMDLATYINMTWSAPLOVITALLYLMLNGPSVLGAVAVMYLMEVFN 480

Db 421 GEIVNLSVDAQRPMDLATYINMTWSAPLOVITALLYLMLNGPSVLGAVAVMYLMEVFN 480

QY 481 AVMAKTKTYOVAMKSKDNRIKLMNELINGIKVLAKIYAMELARKDKYLAIRQELKYLK 540

Db 481 AVMAKTKTYOVAMKSKDNRIKLMNELINGIKVLAKIYAMELARKDKYLAIRQELKYLK 540

QY 541 KSAVLSAVGTFPTWCTPPLVALCTPAVYVTTDENNIIDAOAPFSLAFNLIRPLNLLP 600

Db 541 KSAVLSAVGTFPTWCTPPLVALCTPAVYVTTDENNIIDAOAPFSLAFNLIRPLNLLP 600

QY 601 MVISIVQASVSLKRLRIFLSHEELPDSIRRPVKOGGNTSITVRNATFTWASDPT 660

Db 601 MVISIVQASVSLKRLRIFLSHEELPDSIRRPVKOGGNTSITVRNATFTWASDPT 660

QY 661 LNTGTFSTPEBGLAVAVGVQVCGKSSLSALLAEMDKYEGVAIKGSVAVYPOQAMIOND 720

Db 661 LNTGTFSTPEBGLAVAVGVQVCGKSSLSALLAEMDKYEGVAIKGSVAVYPOQAMIOND 720

QY 721 SLRENILFSGCLEPFRYSVIOACALPDLITLPSGDTETGEGVNLSSGGOKOVSLAR 780

Db 721 SLRENILFSGCLEPFRYSVIOACALPDLITLPSGDTETGEGVNLSSGGOKOVSLAR 780

```
QY 781 AVYSNADIYLFDDPLSAVDAAHGKHI FENVIGPKMKLNKTRILVTSHMSYL PQVDYIIV 840
Db 781 AVYSNADIYLFDDPLSAVDAAHGKHI FENVIGPKMKLNKTRILVTSHMSYL PQVDYIIV 840
QY 841 MSGGKISMGSYQELLARDGAFAEFLRTYASTEQDABENGVTGVS GPGEAKOMENGM 900
Db 841 MSGGKISMGSYQELLARDGAFAEFLRTYASTEQDABENGVTGVS GPGEAKOMENGM 900
QY 901 LVTSAGKOLOROLSSSSSSSGDISRRHNSTAELOKAKAKKEETMKLEADAKAOTGQVKL 960
Db 901 LVTSAGKOLOROLSSSSSSSGDISRRHNSTAELOKAKAKKEETMKLEADAKAOTGQVKL 960
QY 961 SVYDYMKAIGLFTISFLSIPLFMCHNVSALASNYMLSMITDDPIVNGTOEHTKRLSYG 1020
Db 961 SVYDYMKAIGLFTISFLSIPLFMCHNVSALASNYMLSMITDDPIVNGTOEHTKRLSYG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGGILASRCLHVDLHLSILRSPMSFFERTPSGNLVNRFSEL 1080
Db 1021 ALGISOGIAVFGYSMAVSIIGGILASRCLHVDLHLSILRSPMSFFERTPSGNLVNRFSEL 1080
QY 1081 DTVDSMIPEVIMKFMGSLFENVIGACTIVLLATPIAIIIPPLGLIYFFVORFVYASSROL 1140
Db 1081 DTVDSMIPEVIMKFMGSLFENVIGACTIVLLATPIAIIIPPLGLIYFFVORFVYASSROL 1140
QY 1141 KRLESVSRSPYVSHNETLIGSVVIRAFEEQERFIHQSDLKVDENOKAYPSIYANRWLA 1200
Db 1141 KRLESVSRSPYVSHNETLIGSVVIRAFEEQERFIHQSDLKVDENOKAYPSIYANRWLA 1200
QY 1201 VRLCEVNGCIVLPAALFAVIRSHSISAGLVGLSVSLSQVTTYLNMVLRMSSEMETNIVA 1260
Db 1201 VRLCEVNGCIVLPAALFAVIRSHSISAGLVGLSVSLSQVTTYLNMVLRMSSEMETNIVA 1260
QY 1261 VERKEVSETEKEAPWQIOETAPRPSWPQVGRVERRANCCLAYRELDVLAHINVTINGG 1320
Db 1261 VERKEVSETEKEAPWQIOETAPRPSWPQVGRVERRANCCLAYRELDVLAHINVTINGG 1320
QY 1321 EKVGIIVGRTGAGKSLITGLFRINESAGEIIIDGINAKI GLHDLRFKTIIPODPVLF 1380
Db 1321 EKVGIIVGRTGAGKSLITGLFRINESAGEIIIDGINAKI GLHDLRFKTIIPODPVLF 1380
QY 1381 SGLSRMNLDPFSQYSDEEVTWSLELAHLKDFVSALPDKLDHECAEGENLSVGORQLVCL 1440
Db 1381 SGLSRMNLDPFSQYSDEEVTWSLELAHLKDFVSALPDKLDHECAEGENLSVGORQLVCL 1440
QY 1441 ARALLRKTKIIVDEATAVULETDDLIQSTIRPOFECTVLTIAHRLNTIMDTRYIVL 1500
Db 1441 ARALLRKTKIIVDEATAVULETDDLIQSTIRPOFECTVLTIAHRLNTIMDTRYIVL 1500
QY 1501 DKGEIOEGAPSDLLQQRGLFYMAKADAGLV 1531
Db 1501 DKGEIOEGAPSDLLQQRGLFYMAKADAGLV 1531

RESULT 7
AAW9894
ID AAW9894 standard; protein; 1531 AA.
XX
AC AAW9894;
XX
DT 10-JUN-1999 (first entry)
XX
DE Human multidrug resistance-associated protein natural variant.
XX
KM Human; multidrug resistance-associated protein; MRP; cytotoxic drug;
XX
KW cancer; chemotherapy.
XX
OS Homo sapiens.
XX
PN US5891724-A.
XX
PD 06-APR-1999.
XX
PF 05-JUN-1995; 95US-00460907.
```

```
XX 27-OCT-1992; 92US-00966923.
PR 08-MAR-1993; 93US-00029340.
PR 26-OCT-1993; 93US-00141893.
PR 20-MAR-1995; 95US-00407207.
XX
PA (TOOH ) UNIV QUEBENS KINGSTON.
XX
PI Cole SPC, Deeley RG;
XX
DR WPI; 1999-253868/21.
DR N-PSDB; AAX19818.
XX
PT Protecting mammalian cells against cytotoxic drugs.
PS Claim 2; Col 79-86; 82pp; English.
XX
CC The present sequence represents a human multidrug resistance-associated
CC protein (MRP). The present invention also describes a method for
CC protecting a mammalian cell against the cytotoxicity of anthracyclines,
CC epipodophyllotoxins and vinca alkaloids (A) by introducing into it a nucleic
CC acid (II) that encodes a MRP protein (III). Introduction of (I) protects
CC cells against cytotoxic effects of (A), particularly to protect normal
CC cells against (A) being used for treatment of cancers. Cells transformed
CC with (II) can be used to screen for agents that affect multidrug
CC resistance or are directly toxic to multidrug resistant cells, i.e.
CC potential therapeutics for multidrug-resistant cancers. Confering
CC resistance to normal cells should allow an increase in the dose of (A)
CC that can be administered safely
XX
SQ Sequence 1531 AA;
XX
Query Match 79.4%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSADGSDPLMDMNTWNTSNPDFTKCPONTLVWVPCPYLACPPFYLYSRH 60
Db 1 MALRGFCSADGSDPLMDMNTWNTSNPDFTKCPONTLVWVPCPYLACPPFYLYSRH 60
QY 61 DRGYIQMTPLNKTTALGFLMIYCMADLFYSFERSRGIFLAPVFLVSPILGITTLLA 120
Db 61 DRGYIQMTPLNKTTALGFLMIYCMADLFYSFERSRGIFLAPVFLVSPILGITTLLA 120
QY 121 TFLIQLRERRKGVSSGIMLTFWVALVCAIILRSKIMTALKEDAQVDLFRDITFYVYFS 180
Db 121 TFLIQLRERRKGVSSGIMLTFWVALVCAIILRSKIMTALKEDAQVDLFRDITFYVYFS 180
QY 181 LLLIQVLSCGSDRSPLFSETIHDNPNCPRESSASFLSRITFWMTGLIVGRPLGSD 240
Db 181 LLLIQVLSCGSDRSPLFSETIHDNPNCPRESSASFLSRITFWMTGLIVGRPLGSD 240
QY 241 LMSLNKEDTSEOVVPLVVKMKKECAKTRKOPVYVSSKDPAPQPKSSKYDANEVEAL 300
Db 241 LMSLNKEDTSEOVVPLVVKMKKECAKTRKOPVYVSSKDPAPQPKSSKYDANEVEAL 300
QY 301 IVKSPQKEMNPSEKVLKTFGPYFLMSFFPKAIHDLMSGPIILKLILKFVNDTRAPD 360
Db 301 IVKSPQKEMNPSEKVLKTFGPYFLMSFFPKAIHDLMSGPIILKLILKFVNDTRAPD 360
QY 361 WQGYFYVLLFVTRACLOTVLHQVFHICFVSGMRIRKTAIVGAYRRKALVITNSARKSSTV 420
Db 361 WQGYFYVLLFVTRACLOTVLHQVFHICFVSGMRIRKTAIVGAYRRKALVITNSARKSSTV 420
QY 421 GEIVNLSVDAORFMDLATYINMIWSAPLGQIATLYLMLNLPSPVLGAVAVMWLMPVN 480
Db 421 GEIVNLSVDAORFMDLATYINMIWSAPLGQIATLYLMLNLPSPVLGAVAVMWLMPVN 480
QY 481 AVNAAKTKTYQVAHMSKDNRIKLMEILNGIRVILKIYAMELAPKDKVLAIRQELKVLK 540
Db 481 AVNAAKTKTYQVAHMSKDNRIKLMEILNGIRVILKIYAMELAPKDKVLAIRQELKVLK 540
```

```

QY 541 KSAIYISANGTFTWCTPPLVALCTPAVYTTIDENNIIDAQTAFAVSLAFLNIRPPLNLP 600
Db 541 KSAIYISANGTFTWCTPPLVALCTPAVYTTIDENNIIDAQTAFAVSLAFLNIRPPLNLP 600
QY 601 MVISSIVASYSKLKRLIFLSHEELPDSIERRPVKDGGTNSITVRNATFTWASDDPT 660
Db 601 MVISSIVASYSKLKRLIFLSHEELPDSIERRPVKDGGTNSITVRNATFTWASDDPT 660
QY 661 LINGITFSIPEGALVAVGVGVGCKSSLSALLAENDKYEGHVAIKGSVAVYVPOQAMIOND 720
Db 661 LINGITFSIPEGALVAVGVGVGCKSSLSALLAENDKYEGHVAIKGSVAVYVPOQAMIOND 720
QY 721 SLRENITLFGCQLEPEYRSVTIOACALLPDLLEILPESGDRTEIGEKGVNISGGQKQVSLAR 780
Db 721 SLRENITLFGCQLEPEYRSVTIOACALLPDLLEILPESGDRTEIGEKGVNISGGQKQVSLAR 780
QY 781 AVYSNADIIYLPDDPLSAVDAAHAKHIFENVIGPKMKLNKRILVTHSMYSYLPQVDVILV 840
Db 781 AVYSNADIIYLPDDPLSAVDAAHAKHIFENVIGPKMKLNKRILVTHSMYSYLPQVDVILV 840
QY 841 MSGKISMGYSYOEILLARDGAFAEFLRTYASTEOBDAENGVTGSGPGEAKOMENG 900
Db 841 MSGKISMGYSYOEILLARDGAFAEFLRTYASTEOBDAENGVTGSGPGEAKOMENG 900
QY 901 LVTTDSAGQOLQROLSSSSYSGDISRHNSTAELQAKAKKEETWKLMEADKAQTGOYVL 960
Db 901 LVTTDSAGQOLQROLSSSSYSGDISRHNSTAELQAKAKKEETWKLMEADKAQTGOYVL 960
QY 961 SYVMYMAKIGLFIPLSIFLPMCNHVASLANSYMLMTDPIINGQOEHKVALSYVG 1020
Db 961 SYVMYMAKIGLFIPLSIFLPMCNHVASLANSYMLMTDPIINGQOEHKVALSYVG 1020
QY 1021 ALGISQGIAPVFGYSNAVSIIGIILASRCLHVDLHSLILSPMSFFERTPSGNLVNRFSEL 1080
Db 1021 ALGISQGIAPVFGYSNAVSIIGIILASRCLHVDLHSLILSPMSFFERTPSGNLVNRFSEL 1080
QY 1081 DTVDSMITEBVIKMFNGSLFNVTGACIVILLATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
Db 1081 DTVDSMITEBVIKMFNGSLFNVTGACIVILLATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
QY 1141 KRLSEVSRSPLYSHENETLLAGSVIRAFPEQERFIHOSDLKYDENOKAYYPSIVANRWLA 1200
Db 1141 KRLSEVSRSPLYSHENETLLAGSVIRAFPEQERFIHOSDLKYDENOKAYYPSIVANRWLA 1200
QY 1201 VRLSECVGNCIVLFAALFAVISIRHSLSAGLVGSLVSYSIQVTTYLNMLVMSSEMETNIVA 1260
Db 1201 VRLSECVGNCIVLFAALFAVISIRHSLSAGLVGSLVSYSIQVTTYLNMLVMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVFRNYCLARYREDLDPYLRIHINVTINGG 1320
Db 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVFRNYCLARYREDLDPYLRIHINVTINGG 1320
QY 1321 EKVGI VGRGTGAGKSLTGLPRINSABGEIITIDGINIAKIGLHDLRFKITTIIIPDDPVLF 1380
Db 1321 EKVGI VGRGTGAGKSLTGLPRINSABGEIITIDGINIAKIGLHDLRFKITTIIIPDDPVLF 1380
QY 1381 SGLSLMNLDPFSQYDDEEWTSLLELAHLKDFVSALPDKLDHCAEGENLSYGOQOLVCL 1440
Db 1381 SGLSLMNLDPFSQYDDEEWTSLLELAHLKDFVSALPDKLDHCAEGENLSYGOQOLVCL 1440
QY 1441 ARALIRKTIIVLDEATAVADLETDDLQSTIRTOFEDCTVLTIAHRLNTIMDYRIVYL 1500
Db 1441 ARALIRKTIIVLDEATAVADLETDDLQSTIRTOFEDCTVLTIAHRLNTIMDYRIVYL 1500
QY 1501 DKGEIOEGAPSDLLQORGLFYSMAKDAGLV 1531
Db 1501 DKGEIOEGAPSDLLQORGLFYSMAKDAGLV 1531

```

```

AC AA55799;
XX 28-FEB-2000 (first entry)
XX
DE Human multidrug resistance-associated protein (MRP) variant.
XX
KM Chemosensitizer; multidrug resistance-associated protein; MRP; human;
XX therapeutic agent; P-glycoprotein-mediated multidrug resistance; lung;
XX cancer; variant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 685
FT /label= L685S
FT /note= "wild-type Leu is replaced by Ser"
FT
FT Misc-difference 1282
FT /label= R1282A
FT /note= "wild-type Arg is replaced by Ala"
FT
XX
XX US6001563-A.
XX
XX 14-DEC-1999.
XX
XX 05-JUN-1995; 95US-00463179.
XX
XX 27-OCT-1992; 92US-00966923.
XX 08-MAR-1993; 93US-00029340.
XX 26-OCT-1993; 93US-00141893.
XX 20-MAR-1995; 95US-00407207.
XX
XX (TOOH ) UNIV QUEBENS KINGSTON.
XX
XX Cole SP, Deeley RG;
XX
XX WPI: 2000-061877/05.
XX N-PsDB; AA239556.
XX
XX Identification of chemosensitizers useful for treating cancer, using
XX nucleic acids encoding multidrug resistance-associated protein.
XX
XX Claim 3; Col 69-80; 77pp; English.
XX
XX The invention provides a method for identifying a substance which is a
XX chemosensitizer that comprises, contacting a cell transfected with
XX nucleic acid encoding multidrug resistance-associated protein (MRP) with
XX a therapeutic agent in vitro. The method is useful for identifying
XX chemosensitizers which may then be used to treat cancer (especially lung
XX cancer). The method allows the identification of chemosensitizers which
XX do not reverse P-glycoprotein-mediated multidrug resistance. The present
XX sequence represents a human MRP variant
XX
SQ Sequence 1531 AA:
Query Match 79.4%; Score 7860; DB 3; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRFGCADSDPLMDMNNVMTNSNPDTFCQFQNTVYVWPCFYLMACFPPEYLYLSRH 60
Db 1 MALRFGCADSDPLMDMNNVMTNSNPDTFCQFQNTVYVWPCFYLMACFPPEYLYLSRH 60
QY 61 DRGYQMPUNKTKTALGFLIMVCMADLFYSFMRSRGIFLAPVFLVSPTLGITTLLA 120
Db 61 DRGYQMPUNKTKTALGFLIMVCMADLFYSFMRSRGIFLAPVFLVSPTLGITTLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFWLVALCALILRSKIMTALKEDAQVDLPFDITFYVYFS 180
Db 121 TFLIQLERRKGVSSGIMLTFWLVALCALILRSKIMTALKEDAQVDLPFDITFYVYFS 180
QY 181 LLLIOLVLSGSDSPLFSEFTIHDPNPPSSASGLSTITFTWMTGLIVRGYROPLEGSD 240
Db 181 LLLIOLVLSGSDSPLFSEFTIHDPNPPSSASGLSTITFTWMTGLIVRGYROPLEGSD 240

```

RESULT 8
 AA55799
 ID AA55799 standard; protein; 1531 AA.
 XX

QY 241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVKVVYSSKDPACQESSKYDANEVEAL 300
 DB 241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVKVVYSSKDPACQESSKYDANEVEAL 300
 QY 301 IVKSPQKEMNPSLFKVLYKTFGPYFLMSFPFKAIHDLMMFSGPOLIKLLIKFVNDTKAD 360
 DB 301 IVKSPQKEMNPSLFKVLYKTFGPYFLMSFPFKAIHDLMMFSGPOLIKLLIKFVNDTKAD 360
 QY 361 MOGFYYTTLFTVTLCTLOTLVHOVPHICFVSGMRKTAIVIGAVRKALVITNSARKSTV 420
 DB 361 MOGFYYTTLFTVTLCTLOTLVHOVPHICFVSGMRKTAIVIGAVRKALVITNSARKSTV 420
 QY 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOYLALYLLMLNGPSVLAGVAVMLVNVN 480
 DB 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOYLALYLLMLNGPSVLAGVAVMLVNVN 480
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAMELAFKDLAIROBELKYLK 540
 DB 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAMELAFKDLAIROBELKYLK 540
 QY 541 KSAVLASAVGTFTWCTPPLVALCTPFAVYVITDENNIIDAOFAVSLALFNILRPPLNLP 600
 DB 541 KSAVLASAVGTFTWCTPPLVALCTPFAVYVITDENNIIDAOFAVSLALFNILRPPLNLP 600
 QY 601 MVISIYQASVSLKRLRIFLSHEELEPDSIERRPVKGCGNISTVRNATFTMARSDPT 660
 DB 601 MVISIYQASVSLKRLRIFLSHEELEPDSIERRPVKGCGNISTVRNATFTMARSDPT 660
 QY 661 LMGITFSIPGALVAVVQVCGSKSSLSALLAEMDKVEGVAIKGSVAVYPOQAMQND 720
 DB 661 LMGITFSIPGALVAVVQVCGSKSSLSALLAEMDKVEGVAIKGSVAVYPOQAMQND 720
 QY 721 SLRENIIIFGCOLLEPPYRSVIOACALLPDLIILPSGDRTEIGEKNVLSGGQKORVSLAR 780
 DB 721 SLRENIIIFGCOLLEPPYRSVIOACALLPDLIILPSGDRTEIGEKNVLSGGQKORVSLAR 780
 QY 781 AVVSNADIYLEDDBLSAVDAHVGHIFENVIIGPKGMKNKRIIIVTSMSTLPQVDYVITV 840
 DB 781 AVVSNADIYLEDDBLSAVDAHVGHIFENVIIGPKGMKNKRIIIVTSMSTLPQVDYVITV 840
 QY 841 MSGGKISEMGSYQELIARDAFAEFLFTYASTEOEODAEENGVTGSGPGKEAKOMENG 900
 DB 841 MSGGKISEMGSYQELIARDAFAEFLFTYASTEOEODAEENGVTGSGPGKEAKOMENG 900
 QY 901 LVYDSAGKOLROUJSSSSSYSGDISRHHNSTAEIQAQAKKEETWKLMEADKAQTQVYKL 960
 DB 901 LVYDSAGKOLROUJSSSSSYSGDISRHHNSTAEIQAQAKKEETWKLMEADKAQTQVYKL 960
 QY 961 SVYVDYMKATGLFTSPISITFLMGNHVSALASNWLSLMTDDPIVNGOEHTRKRLSVYG 1020
 DB 961 SVYVDYMKATGLFTSPISITFLMGNHVSALASNWLSLMTDDPIVNGOEHTRKRLSVYG 1020
 QY 1021 ALGISOGIAVFGVMAVSIIGIILASRCLHVDLHLSILSPMSFEPTPSGULVNRFSKEL 1080
 DB 1021 ALGISOGIAVFGVMAVSIIGIILASRCLHVDLHLSILSPMSFEPTPSGULVNRFSKEL 1080
 QY 1081 DTVDMSIPEVIKMFMSGLFNVIACIYILLATPIAIIIPPLGIYEFVQGFYVASSROL 1140
 DB 1081 DTVDMSIPEVIKMFMSGLFNVIACIYILLATPIAIIIPPLGIYEFVQGFYVASSROL 1140
 QY 1141 KRLESVRSRPSYSHFNFTLGVSVIRAFEBEERFIHOSDLKVDENQRAYPSIYANWMLA 1200
 DB 1141 KRLESVRSRPSYSHFNFTLGVSVIRAFEBEERFIHOSDLKVDENQRAYPSIYANWMLA 1200
 QY 1201 VRLIECVNGCIYLFPAALFAVISRHSLSAGLVGLSVYSLSQVTTYLNLVMSSEMETNIVA 1260
 DB 1201 VRLIECVNGCIYLFPAALFAVISRHSLSAGLVGLSVYSLSQVTTYLNLVMSSEMETNIVA 1260
 QY 1261 VERLKEVSETEKAPMOIOETAPSSWPQVGRVFRNYCLARYRDLDPVLRIHIVNTINGG 1320
 DB 1261 VERLKEVSETEKAPMOIOETAPSSWPQVGRVFRNYCLARYRDLDPVLRIHIVNTINGG 1320
 QY 1320 VERLKEVSETEKAPMOIOETAPSSWPQVGRVFRNYCLARYRDLDPVLRIHIVNTINGG 1320
 DB 1320 VERLKEVSETEKAPMOIOETAPSSWPQVGRVFRNYCLARYRDLDPVLRIHIVNTINGG 1320

QY 1321 EKVGIVERTGAKSSLTGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPYLF 1380
 DB 1321 EKVGIVERTGAKSSLTGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPYLF 1380
 QY 1381 SGSLRNMULDPPSQYSDEEWTSLELAHLKDFVSALPKLDHECAEGGENISVQGRQIVCL 1440
 DB 1381 SGSLRNMULDPPSQYSDEEWTSLELAHLKDFVSALPKLDHECAEGGENISVQGRQIVCL 1440
 QY 1441 ABALLRRTKTLIVDEATRAVDLETFDDLIQSTIRFOFEDCTVLTAAHRLNTMDYTRVYL 1500
 DB 1441 ABALLRRTKTLIVDEATRAVDLETFDDLIQSTIRFOFEDCTVLTAAHRLNTMDYTRVYL 1500
 QY 1501 DKGEIQEYGA PSDLLQORGLFYSMAKXAGLY 1531
 DB 1501 DKGEIQEYGA PSDLLQORGLFYSMAKXAGLY 1531
 RESULT 9
 ID AAY78873 standard; protein: 1531 AA.
 XX AAY78873;
 AC AAY78873;
 XX 19-MAY-2000 (first entry)
 DT 19-MAY-2000 (first entry)
 XX
 DE Multidrug resistance protein (MRP) natural variant amino acid sequence.
 XX
 KM Multidrug resistance protein; MRP; human; anthracycline; Vinca alkaloid;
 KM epipodophyllotoxin; cancer; leukaemia.
 XX Homo sapiens.
 OS
 XX US6025473-A.
 PN 15-FEB-2000.
 PD 15-FEB-2000.
 XX
 PF 05-JUN-1995; 95US-00461384.
 XX
 PR 27-OCT-1992; 92US-00966923.
 PR 08-MAR-1993; 93US-00029340.
 PR 26-OCT-1993; 93US-00141893.
 PR 20-MAR-1995; 95US-00407207.
 XX
 PA (TOOH) UNIV QUEENS KINGSTON.
 XX
 PI Cole SPC, Deeley RG;
 PI WPI; 2000-181838/16.
 XX N-PSDB; AAZ90193.
 DR
 PT Isolated protein conferring multidrug resistance, to at least two drugs
 PT selected from anthracyclines, epipodophyllotoxins and Vinca alkaloids, on
 PT a drug sensitive mammalian cell.
 XX
 PS Claim 10; Col 79-88; 78pp; English.
 XX
 CC This sequence represents a human multidrug resistance protein (MRP)
 CC natural variant amino acid sequence. The human MRP confers multidrug
 CC resistance, including resistance to at least two drugs selected from
 CC anthracyclines, epipodophyllotoxins and Vinca alkaloids, on a drug
 CC sensitive mammalian cell, when the protein is expressed in the cell. The
 CC multidrug resistance is not substantially reversed by chemosensitizers
 CC which reverse P-glycoprotein-mediated multidrug resistance. The MRP
 CC protein sequence can be used to generate antibodies against MRP. The MRP
 CC protein and nucleotide sequences can be used in compositions which are
 CC used to treat patients with tumours displaying multidrug resistance. The
 CC compositions and methods of the invention can be used particularly to
 CC treat breast cancer, leukaemias, fibrosarcomas, cervical cancer, gliomas,
 CC thymomas, neuroblastomas, and lung cancer. Antibodies directed against
 CC MRP can be used to inhibit the multidrug resistance of a multidrug
 CC resistant cell
 XX
 SO Sequence 1531 AA;

Query Match	79.4%; Score 7860; DB 3; Length 1531;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1531; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRFGSADSDSDLMNNVTMTNSPDKFCQONTVYVWPCFTLMACFEPFLYLSRH 60	
DB 1 MALRFGSADSDSDLMNNVTMTNSPDKFCQONTVYVWPCFTLMACFEPFLYLSRH 60	
QY 61 DRGIYQMTPLNKTKTALGFLIMIVCMADLFYSFWERSRGIFLAPFLVSPILLGITTLA 120	
DB 61 DRGIYQMTPLNKTKTALGFLIMIVCMADLFYSFWERSRGIFLAPFLVSPILLGITTLA 120	
QY 121 TFLIQLEBRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVFS 180	
DB 121 TFLIQLEBRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVFS 180	
QY 181 LLLIQVLVSCSDSPLEFSETIHDNPPCESASLSTITFMWITGLIVRGYRQLEBSD 240	
DB 181 LLLIQVLVSCSDSPLEFSETIHDNPPCESASLSTITFMWITGLIVRGYRQLEBSD 240	
QY 241 LMSLNKEDTSEQVAVVLVKNMKKECAKTRKQPVKVYSSKDPACPKESSKYDANEVVAL 300	
DB 241 LMSLNKEDTSEQVAVVLVKNMKKECAKTRKQPVKVYSSKDPACPKESSKYDANEVVAL 300	
QY 301 IVKSPQKEMNPSLFKVLKTFGPYELMSFFPKAHLDMFSGPQILKLLKRVNDTKAPD 360	
DB 301 IVKSPQKEMNPSLFKVLKTFGPYELMSFFPKAHLDMFSGPQILKLLKRVNDTKAPD 360	
QY 361 WQGFYTYLLFVTACLOTLVLAHQYHICFVSGMRKTAIVIGAVYKALVITNSAKSSTV 420	
DB 361 WQGFYTYLLFVTACLOTLVLAHQYHICFVSGMRKTAIVIGAVYKALVITNSAKSSTV 420	
QY 421 GEIVNLMSVDKORFMDLATYINMISAPLOVILALYLMLNIGPSVLGAVAVMLVAVN 480	
DB 421 GEIVNLMSVDKORFMDLATYINMISAPLOVILALYLMLNIGPSVLGAVAVMLVAVN 480	
QY 481 AVMAKTKTYQVAHMKSKONRIKLMNEILNGIKVLKLYAMBLAFKQYLAIROBELKYLK 540	
DB 481 AVMAKTKTYQVAHMKSKONRIKLMNEILNGIKVLKLYAMBLAFKQYLAIROBELKYLK 540	
QY 541 KSAIYLSANGTPTWCTPFLVALCTPAVYVTTIDENNILDAQTAFFVSLAFNLIIRPLNLI 600	
DB 541 KSAIYLSANGTPTWCTPFLVALCTPAVYVTTIDENNILDAQTAFFVSLAFNLIIRPLNLI 600	
QY 601 MVISIVQASVSLKRLRIFLSHEBLEPDSIERRPVKDGGGINSITVRNATFTWASDDEPT 660	
DB 601 MVISIVQASVSLKRLRIFLSHEBLEPDSIERRPVKDGGGINSITVRNATFTWASDDEPT 660	
QY 661 LNGITFSIPEGALVAVVGVGGKSSLLSALLAEMDKYEGHVAIKGSVAAYVPOQAMIQND 720	
DB 661 LNGITFSIPEGALVAVVGVGGKSSLLSALLAEMDKYEGHVAIKGSVAAYVPOQAMIQND 720	
QY 721 SLRENIILFSCGLEEYYSVIOACALLPDLLETLPESGDRTEIGEKGVNLSGGCKQVSLAR 780	
DB 721 SLRENIILFSCGLEEYYSVIOACALLPDLLETLPESGDRTEIGEKGVNLSGGCKQVSLAR 780	
QY 781 AVYSNADLYLPDDPLSAVDAAHVKHIFENVIGPKGMKNKRIILVTHSNSTYLPQVDVILV 840	
DB 781 AVYSNADLYLPDDPLSAVDAAHVKHIFENVIGPKGMKNKRIILVTHSNSTYLPQVDVILV 840	
QY 841 MSGKISMGVSVOELLARDGAFAEFLRTYASTEOBDAEENGVTGSGPGEAKOMENGM 900	
DB 841 MSGKISMGVSVOELLARDGAFAEFLRTYASTEOBDAEENGVTGSGPGEAKOMENGM 900	
QY 901 LVTTDSAGQOLRQLSSSSSYSGDISRHNHSTALQKAKEKKEPTWKLMEADRAQOQYVL 960	
DB 901 LVTTDSAGQOLRQLSSSSSYSGDISRHNHSTALQKAKEKKEPTWKLMEADRAQOQYVL 960	
QY 961 SYVMDYMAIGLFIISFLSIFLPMCNHVSALASNWMLSTMTDPIYNGTOEHKXVLSYVG 1020	
DB 961 SYVMDYMAIGLFIISFLSIFLPMCNHVSALASNWMLSTMTDPIYNGTOEHKXVLSYVG 1020	

QY 1021 ALGISGIAVFGYSMAVSIIGIILASRCLHYDHLHSILRSPMSFERPTPSGULVNRFSKEL 1080	
DB 1021 ALGISGIAVFGYSMAVSIIGIILASRCLHYDHLHSILRSPMSFERPTPSGULVNRFSKEL 1080	
QY 1081 DTVDMSIPEVYKMGMSLFENVIGACIYILLATPIAIIIPPLGIYFFVQGFYVASSROL 1140	
DB 1081 DTVDMSIPEVYKMGMSLFENVIGACIYILLATPIAIIIPPLGIYFFVQGFYVASSROL 1140	
QY 1141 KRLESVSRSPYSHFNETLLGVSVIRAFEEQERPIHOSDLKVDENOKAYYPSIVANRWLA 1200	
DB 1141 KRLESVSRSPYSHFNETLLGVSVIRAFEEQERPIHOSDLKVDENOKAYYPSIVANRWLA 1200	
QY 1201 VRLCEVGNCTYLPALPAVVISRHSLSAGLVLSYSLOVTTYANMLVRMSSEMETNIVA 1260	
DB 1201 VRLCEVGNCTYLPALPAVVISRHSLSAGLVLSYSLOVTTYANMLVRMSSEMETNIVA 1260	
QY 1261 VERLKEYSETEKEAPMOIOETAPPSMPQYGRVFEFRNYCLRYRBDLPVLHINVTINGG 1320	
DB 1261 VERLKEYSETEKEAPMOIOETAPPSMPQYGRVFEFRNYCLRYRBDLPVLHINVTINGG 1320	
QY 1321 EKVGIVGRGTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRPKITIIPODPVLF 1380	
DB 1321 EKVGIVGRGTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRPKITIIPODPVLF 1380	
QY 1381 SCSLRNMLDPPSQYSDERWTSLELAHLKDFVSLPDKLDEHCAEGGENLSVGROLVCL 1440	
DB 1381 SCSLRNMLDPPSQYSDERWTSLELAHLKDFVSLPDKLDEHCAEGGENLSVGROLVCL 1440	
QY 1441 ARALLRKTILVLEATPAVVDLETTDDLIOSTIRFOPEDCTVLTIAHRLNTIMDYTRVIL 1500	
DB 1441 ARALLRKTILVLEATPAVVDLETTDDLIOSTIRFOPEDCTVLTIAHRLNTIMDYTRVIL 1500	
QY 1501 DKGEIOEYGAPSDLLQQRGLFYSNAKDAGLV 1531	
DB 1501 DKGEIOEYGAPSDLLQQRGLFYSNAKDAGLV 1531	
RESULT 10	
ABG61810	
ID ABG61810 standard; protein; 1531 AA.	
XX	
AC ABG61810;	
XX	
DT 15-AUG-2002 (first entry)	
XX	
DE Prostate cancer-associated protein #11.	
XX	
KM Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.	
XX	
OS Mammalia.	
XX	
PN WO200230268-A2.	
XX	
PD 18-APR-2002.	
XX	
PF 12-OCT-2001; 2001WO-US032045.	
XX	
PR 13-OCT-2000; 2000US-00687576.	
PR 08-DEC-2000; 2000US-00733288.	
PR 08-DEC-2000; 2000US-00733742.	
PR 24-JAN-2001; 2001US-0263957P.	
PR 16-MAR-2001; 2001US-0276791P.	
PR 16-MAR-2001; 2001US-0276888P.	
PR 06-APR-2001; 2001US-0281922P.	
PR 24-APR-2001; 2001US-0286214P.	
PR 30-APR-2001; 2001US-00847046.	
PR 04-MAY-2001; 2001US-0288589P.	
XX	
PA (BOSB-) EOS BIOTECHNOLOGY INC.	
XX	
PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;	
XX	
DR WPI, 2002-471335/50.	

DR N-PSDB; ABK92125.
XX Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.
PS Claim 27; Page 309; 436pp; English.
XX
XX The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC associating a biological sample from the patient with prostate cancer-
CC associated polynucleotides (designated PC genes) that selectively
CC hybridise to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins
CC
XX
SQ Sequence 1531 AA;
Query Match 79.4%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGPSADSDPLDMNMTNTMTSNPFTGQNTLVLPWCYLMACFPFYLYLSRH 60
DB 1 MALRGPSADSDPLDMNMTNTMTSNPFTGQNTLVLPWCYLMACFPFYLYLSRH 60
QY 61 DRGVIQMTPLNKTALGFLMIYCMADLFYSFWERSGFIAPVFLVPTLIGITLLA 120
DB 61 DRGVIQMTPLNKTALGFLMIYCMADLFYSFWERSGFIAPVFLVPTLIGITLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFWALVLCALILRSKIMTALKEDAOVDLFFDITFYVFS 180
DB 121 TFLIQLERRKGVSSGIMLTFWALVLCALILRSKIMTALKEDAOVDLFFDITFYVFS 180
QY 181 LLLIQLVLSGSDSPISFETIHDNPPCESASFLSITFWMTGLIVRGYRPLGSD 240
DB 181 LLLIQLVLSGSDSPISFETIHDNPPCESASFLSITFWMTGLIVRGYRPLGSD 240
QY 241 LMSLNKEDTSQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPKSSSKVDANEVEAL 300
DB 241 LMSLNKEDTSQVVPVLVKNWKECAKTRKQPVKVYSSKDPAPKSSSKVDANEVEAL 300
QY 301 IVKSPOKEWNPDLFKVLYKTGFPYFLMSFFPKAIDHLMFSGPOLIKLIFVNDTRAPD 360
DB 301 IVKSPOKEWNPDLFKVLYKTGFPYFLMSFFPKAIDHLMFSGPOLIKLIFVNDTRAPD 360
QY 361 WQGFYVVLVPTACIQTLVHOFPHICFVSGMRITKRAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYVVLVPTACIQTLVHOFPHICFVSGMRITKRAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIYNLMSVDAQRFMDLATYINMIWSAPLQVILALYLLMLNGPSVLAVWMLPVN 480
DB 421 GEIYNLMSVDAQRFMDLATYINMIWSAPLQVILALYLLMLNGPSVLAVWMLPVN 480
QY 481 AVNMAKTKTYOVAMKSKDNRIKLMNBIINGIKVLYKLYAMELAFKRVLAIRQELKVLK 540
DB 481 AVNMAKTKTYOVAMKSKDNRIKLMNBIINGIKVLYKLYAMELAFKRVLAIRQELKVLK 540
QY 541 KSAVLAVGFTWCTPFLVALCTFAVYVITDENNNIIDAQAFSLFNLFRPLNLP 600
DB 541 KSAVLAVGFTWCTPFLVALCTFAVYVITDENNNIIDAQAFSLFNLFRPLNLP 600
QY 601 MVISSIVQASVSLKRLIFLSHEELBPSLERRPVKGGGNTSITVNAFTWARSDDPT 660
DB 601 MVISSIVQASVSLKRLIFLSHEELBPSLERRPVKGGGNTSITVNAFTWARSDDPT 660

DB 601 MVISSIVQASVSLKRLIFLSHEELBPSLERRPVKGGGNTSITVNAFTWARSDDPT 660
QY 661 LNTGTTTTPGALVAVVGVCCGSSLLSALLMMDVBEHVAAKGSVAVYPOQAWIQND 720
DB 661 LNTGTTTTPGALVAVVGVCCGSSLLSALLMMDVBEHVAAKGSVAVYPOQAWIQND 720
QY 721 SLRENIIFGCOLBEPYRSYIQAICALPDLEILPSGRTIEGKXNLSGGQKQVSLAR 780
DB 721 SLRENIIFGCOLBEPYRSYIQAICALPDLEILPSGRTIEGKXNLSGGQKQVSLAR 780
QY 781 AVYSNADIVLPDDELSAVDAHVGHKIPENYIGPKMLKNKTRILVTHSMSYLPOVDYIIV 840
DB 781 AVYSNADIVLPDDELSAVDAHVGHKIPENYIGPKMLKNKTRILVTHSMSYLPOVDYIIV 840
QY 841 MSGGKISEMGSYQGLARDDGFAFPLFTYASTEOBDAENGVTVSGPGKEAQMENG 900
DB 841 MSGGKISEMGSYQGLARDDGFAFPLFTYASTEOBDAENGVTVSGPGKEAQMENG 900
QY 901 LVTDASAKOLQROSLSSSSYSGDISRHNSFAELQKAEKKEETWKLMEADKAQTGVKL 960
DB 901 LVTDASAKOLQROSLSSSSYSGDISRHNSFAELQKAEKKEETWKLMEADKAQTGVKL 960
QY 961 SYVWDYMKALGLFISPLIFLMCNHVSALASNTWLSLMTDDPIVNGTOBHTKVLRSYG 1020
DB 961 SYVWDYMKALGLFISPLIFLMCNHVSALASNTWLSLMTDDPIVNGTOBHTKVLRSYG 1020
QY 1021 ALGISQGIAPFGISMAVSIIGIILASRCLAYDLHSILRSFMSFEPFPGNLVNFSEKL 1080
DB 1021 ALGISQGIAPFGISMAVSIIGIILASRCLAYDLHSILRSFMSFEPFPGNLVNFSEKL 1080
QY 1081 DTVDMSIPEVIKPMGSLFNVIACIYILATPAAIIPPLIGIYFVORFYVASSROL 1140
DB 1081 DTVDMSIPEVIKPMGSLFNVIACIYILATPAAIIPPLIGIYFVORFYVASSROL 1140
QY 1141 KRLESVRSFVYSHFNFTLGVSVIRAFEEQERFIHOSDLKVDENQKAYYPSIVANWLA 1200
DB 1141 KRLESVRSFVYSHFNFTLGVSVIRAFEEQERFIHOSDLKVDENQKAYYPSIVANWLA 1200
QY 1201 VRLBVCNCTVLPALFAVISRHSLSAGLVGSYSVLQVTTYINMLVRMSSEMETNIVA 1260
DB 1201 VRLBVCNCTVLPALFAVISRHSLSAGLVGSYSVLQVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSEMEKAPMOIOTAPSSWPQVGRVFRNYCLRYREDLDFVLRHINVTING 1320
DB 1261 VERLKEYSEMEKAPMOIOTAPSSWPQVGRVFRNYCLRYREDLDFVLRHINVTING 1320
QY 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SSSIIRNMLDPPSOYSDEEVTSLBLAHLKDFVSALPKLDHECAGEGENTSVGQROLVCL 1440
DB 1381 SSSIIRNMLDPPSOYSDEEVTSLBLAHLKDFVSALPKLDHECAGEGENTSVGQROLVCL 1440
QY 1441 ARALLRRTKTLVDEAAVADLETDLOSTIRQFEDCTVLTIAHLNLTIMOTRYIVL 1500
DB 1441 ARALLRRTKTLVDEAAVADLETDLOSTIRQFEDCTVLTIAHLNLTIMOTRYIVL 1500
QY 1501 DKGEIOEYGA PSDLQOQGLFYSGMAKDAGLV 1531
DB 1501 DKGEIOEYGA PSDLQOQGLFYSGMAKDAGLV 1531
RESULT 11
ID ABM35012 standard; protein; 1531 AA.
XX ABM35012;
XX AC
XX DT 08-OCT-2003 (first entry)
XX Cancer based on CYP3A5 related protein SEQ ID NO:678.
XX

KW	Cancer; CYP3A5; irinotecan; pharmaceutical; malignant glioma;
KW	cytochrome p450; subfamily 11A; nifedipine oxidase; polypeptide 5;
KW	cytostatic.
XX	
OS	Unidentified.
XX	
PN	WO2003013534-A2.
XX	
PD	20-FEB-2003.
XX	
PF	23-JUL-2002; 2002WO-EP008219.
XX	
PR	23-JUL-2001; 2001EP-00117608.
PR	24-MAY-2002; 2002EP-00011710.
XX	
PA	(EPID-) EPIDAUROS BIOTECHNOLOGIE AG.

PI	Heinrich G, Kerb R;
XX	
DR	WPI; 2003-26814/26.

PT New use of nitroceftoxime for preparation of compositions for treating cancer
PT in subject having genome with variant allele comprising cytochrome p450,
PT subfamily 11A, polypeptide 5 polynucleotide, termed CYP3A5.

PS Disclosure; SEQ ID NO 678; 86pp; English.

CC The present invention describes the use of trinitectan (I) or its
CC derivative for the preparation of a pharmaceutical composition for
CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic
CC cancer, or malignant glioma in a subject having a genome with a variant
CC allele which comprises a cytochrome p450, subfamily IIA (nifedipine
CC oxidase), polypeptide 5 (CYP3A5) polymucleotide (II). (I) and (II) have
CC cytostatic activity. The therapeutic applications of (I) is improved,
CC since it is possible to individually treat a subject with an appropriate
CC dosage and/or an appropriate derivative of (I). Therefore, undesirable,
CC harmful or toxic effects are efficiently avoided. Unnecessary and
CC potentially harmful treatment of those subjects who do not respond to the
CC treatment with substances (nonresponders), as well as the development of
CC drug resistances due to suboptimal drug dosing can be avoided. ACF62200
CC to ACF62751 and ABM34912 to ABM35013 represent sequences used in the
CC exemplification of the present invention

Sequence 1531 AA;
SQ

Query Match	79.4%	Score	7860	DB	6	Length	1531
Best Local Similarity	100.0%	Pred. No.	0				
Matches 1531; Conservative	0	Mismatches	0	Indels	0	Gaps	0

QY	1	MALRFGCSADGSDPLMDNMVVTWNTSNBPTKCGQNVLVWVPEFYIMACPPFFYLTSR	60
Db	1	MALRFGCSADGSDPLMDNMVVTWNTSNBPTKCGQNVLVWVPEFYIMACPPFFYLTSR	60
QY	61	DRGYIOMTPINTKTAALGFLMTVCADL.FYSWERSRGIPLAPVFLVSTLLGITYLLA	120
Db	61	DRGYIOMTPINTKTAALGFLMTVCADL.FYSWERSRGIPLAPVFLVSTLLGITYLLA	120
QY	121	TFLLIOLERRKGVOSSGIMLTFMWLVAALVCAALAIRSKIMTALKEADAQVDFRDIITFYVYS	180
Db	121	TFLLIOLERRKGVOSSGIMLTFMWLVAALVCAALAIRSKIMTALKEADAQVDFRDIITFYVYS	180
QY	181	LLLILOLVLSCFSDRSPFLFSETIHDPNPCBSSASPLSRITFWMTGLIYRGYQPIEGSD	240
Db	181	LLLILOLVLSCFSDRSPFLFSETIHDPNPCBSSASPLSRITFWMTGLIYRGYQPIEGSD	240
QY	241	LMSLNKEDTSEQOVVPVLVNMWKECAKTRQPVKYVSSKDPAQPKSSKVDANEVEAL	300
Db	241	LMSLNKEDTSEQOVVPVLVNMWKECAKTRQPVKYVSSKDPAQPKSSKVDANEVEAL	300
QY	301	IVKSPQKKNPBLFKVLKXTGPGYFLMSPPFKAIHILMFSGQIILKLIFVNDTRKAP	360
Db	301	IVKSPQKKNPBLFKVLKXTGPGYFLMSPPFKAIHILMFSGQIILKLIFVNDTRKAP	360

QY	361	WQGFYTVLLFWA	QJTVLHQQYFHF	CFUSGMRITAVI	GAAYRKAIVTNSAR	SSV	420								
Db	361	WQGFYTVLLFWA	QJTVLHQQYFHF	CFUSGMRITAVI	GAAYRKAIVTNSAR	SSV	420								
QY	421	GEIYNLMSVDA	QFMDATYINNI	WSAPLOVILATY	LMNLGSPVSLAGAVWLV	WPVN	480								
Db	421	GEIYNLMSVDA	QFMDATYINNI	WSAPLOVILATY	LMNLGSPVSLAGAVWLV	WPVN	480								
QY	481	AVMAKRTQY	QVAHMSKDNRI	KLMNEILNGI	KVLKYLAMELA	PKOVTAIROEBELKVLK	540								
Db	481	AVMAKRTQY	QVAHMSKDNRI	KLMNEILNGI	KVLKYLAMELA	PKOVTAIROEBELKVLK	540								
QY	541	KSATLSAVGTF	FWCPEFVVAL	CTFAYVYTI	IDENNILDAQTA	VAVSLAFNIIIRFP	PLNIIIP	600							
Db	541	KSATLSAVGTF	FWCPEFVVAL	CTFAYVYTI	IDENNILDAQTA	VAVSLAFNIIIRFP	PLNIIIP	600							
QY	601	MVSIISYQAS	YSIKRLIIFLS	SHELEPDSIER	RPVKDGGTNSI	YRNATFTMAR	SQDPT	660							
Db	601	MVSIISYQAS	YSIKRLIIFLS	SHELEPDSIER	RPVKDGGTNSI	YRNATFTMAR	SQDPT	660							
QY	661	LNGITFSI	PEGALVAVV	QVCGCKSIL	SALALAE	MDKVEGHVAIKGS	VAVVPOQAMIQND	720							
Db	661	LNGITFSI	PEGALVAVV	QVCGCKSIL	SALALAE	MDKVEGHVAIKGS	VAVVPOQAMIQND	720							
QY	721	SLRBNIIIFG	QLEPPYRSYIQA	CALLPDEIIP	SGDPRITIGRK	GVNLSSGQCOR	VSILAR	780							
Db	721	SLRBNIIIFG	QLEPPYRSYIQA	CALLPDEIIP	SGDPRITIGRK	GVNLSSGQCOR	VSILAR	780							
QY	781	AVYENADIY	FDDPLS	AVDAHWGKIF	PEYVIGPKGM	KNKTRLLYTH	SMSYLPOVDYIY	840							
Db	781	AVYENADIY	FDDPLS	AVDAHWGKIF	PEYVIGPKGM	KNKTRLLYTH	SMSYLPOVDYIY	840							
QY	841	MSGKISEMS	YOELLAR	DGAFAPFL	KTYASTBO	EOBAE	ENGTVGSGPCKEAKOMENG	900							
Db	841	MSGKISEMS	YOELLAR	DGAFAPFL	KTYASTBO	EOBAE	ENGTVGSGPCKEAKOMENG	900							
QY	901	LYTDSACKO	IORULSSSS	SGDIRSH	HNSTAEILQ	RAKKEBTW	KLMENDAKOQOYKL	960							
Db	901	LYTDSACKO	IORULSSSS	SGDIRSH	HNSTAEILQ	RAKKEBTW	KLMENDAKOQOYKL	960							
QY	961	SVYWDYKAI	GLFISFLI	FPCNHN	SALASNYW	LSMTDDPI	VNGTOHTV	RLSVYG	1020						
Db	961	SVYWDYKAI	GLFISFLI	FPCNHN	SALASNYW	LSMTDDPI	VNGTOHTV	RLSVYG	1020						
QY	1021	ALGISOGI	IAFGVSM	AVISGII	ILASRLCH	VDLHSLIR	SPMSFPERP	SGNLVNR	FSKEL	1080					
Db	1021	ALGISOGI	IAFGVSM	AVISGII	ILASRLCH	VDLHSLIR	SPMSFPERP	SGNLVNR	FSKEL	1080					
QY	1081	DTVDSMT	PEVIKOF	MSGLE	FNVI	GCIVIL	ATPIAIIIP	PLGLIYF	FVORFY	VASSROL	1140				
Db	1081	DTVDSMT	PEVIKOF	MSGLE	FNVI	GCIVIL	ATPIAIIIP	PLGLIYF	FVORFY	VASSROL	1140				
QY	1141	KRLSVSR	SPYSHFN	ETLLIG	VSIVTRA	FEEOERFI	HOSDLK	VENDQ	AAVPSI	VANRWLA	1200				
Db	1141	KRLSVSR	SPYSHFN	ETLLIG	VSIVTRA	FEEOERFI	HOSDLK	VENDQ	AAVPSI	VANRWLA	1200				
QY	1201	VRLECVNG	NCIALF	ALPAV	ISRHLS	SGAVGLSV	SYSLQV	TTYLAN	MLVRMS	SEMETNIYA	1260				
Db	1201	VRLECVNG	NCIALF	ALPAV	ISRHLS	SGAVGLSV	SYSLQV	TTYLAN	MLVRMS	SEMETNIYA	1260				
QY	1261	VERLKEY	SETEKEA	PWOIO	ETAP	PSSWPQY	GRVAF	FRNYCL	YREDD	DFVLRI	HNVTINGG	1320			
Db	1261	VERLKEY	SETEKEA	PWOIO	ETAP	PSSWPQY	GRVAF	FRNYCL	YREDD	DFVLRI	HNVTINGG	1320			
QY	1321	EKVGI	VGRTAG	KSSLT	LGLFR	INES	ABGIIID	GINIA	KIGL	HLDR	FKITTIIP	PODPLF	1380		
Db	1321	EKVGI	VGRTAG	KSSLT	LGLFR	INES	ABGIIID	GINIA	KIGL	HLDR	FKITTIIP	PODPLF	1380		
QY	1381	SGSIR	RMULD	PPSOY	SDE	VWTSLE	LALKOF	VSALPK	LDHEC	ABEGEN	LSVGO	ROLVCL	1440		
Db	1381	SGSIR	RMULD	PPSOY	SDE	VWTSLE	LALKOF	VSALPK	LDHEC	ABEGEN	LSVGO	ROLVCL	1440		
QY	1441	ARALLR	TKTILV	DEA	RAV	DLF	DDIIO	STIRI	POF	ED	CVLVI	IAHLANT	IMDYTR	YIVL	1500

Db 1441 ARAIRKTKILVBEATAVDLEFDDLIQSTIRQFEDCTVLTIAHRLNTIMDTYRVL 1500
 QY 1501 DKGEIOEGAPSDLLQQRGLFYMAKQAGLV 1531
 Db 1501 DKGEIOEGAPSDLLQQRGLFYMAKQAGLV 1531

RESULT 12
 ADB20865
 ID ADB20865 standard; protein, 1531 AA.
 AC ADB20865;
 XX 20-NOV-2003 (first entry)
 DE MRP1 based cancer related protein SEQ ID NO:678.
 KM irinotecan; colorectal cancer; cervical cancer; gastric cancer;
 KM lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
 KM variant allele; multidrug resistance protein 1; MRP1; cytosolic.
 OS Unidentified.
 XX WO2003013533-A2.
 XX 20-FEB-2003.
 XX 23-JUL-2002; 2002WO-EP008200.
 XX 23-JUL-2001; 2001EP-00117608.
 PR 24-MAY-2002; 2002EP-00011710.
 PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
 PI Heinrich G, Kerb R;
 DR WPI; 2003-354397/33.

PT Use of irinotecan or its derivative for preparation of a pharmaceutical
 PT composition for treating cancer in a subject having a genome with a
 PT variant allele comprising a multidrug resistance protein 1
 PT polynucleotide.
 XX
 PS Disclosure; SEQ ID NO 678; 100pp; English.
 XX
 CC The present invention describes a method for the use of irinotecan (I) or
 CC its derivative for the preparation of a pharmaceutical composition for
 CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic
 CC cancer, or malignant glioma in a subject having a genome with a variant
 CC allele which comprises a multidrug resistance protein 1 (MRP1)
 CC polynucleotide (II). (I) has cytostatic activity. (I) or its derivative
 CC can be used for the preparation of a pharmaceutical composition for
 CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic
 CC cancer, or malignant glioma in a subject, where the subject is a human
 CC (preferably African or Asian) or a mouse. The present sequence represents
 CC a sequence which is used in the exemplification of the present invention.
 XX
 SQ Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 6; Length 1531;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCADGSDPLMDMNTWNTSNDFKCFONTLVVWPCFYLMACEPFYLYSRH 60
 Db 1 MALRGFCADGSDPLMDMNTWNTSNDFKCFONTLVVWPCFYLMACEPFYLYSRH 60
 QY 61 DRGICQMTPLNKTALGFLIMLYCQADLTFSEMERGRGIFLAPVFLVSPLLGITTLLA 120
 Db 61 DRGICQMTPLNKTALGFLIMLYCQADLTFSEMERGRGIFLAPVFLVSPLLGITTLLA 120
 QY 121 TFLIQLERRRGVSGSGLMFLFWALVCAALILRSKIMTLAKEDAQVDLPFDITFYVYFS 180

Db 121 TFLIQLERRRGVSGSGLMFLFWALVCAALILRSKIMTLAKEDAQVDLPFDITFYVYFS 180
 QY 181 LLLIQLVLSCTSDRSPLPSETIHDNCPSSASFLSRITFWWTGILVRGYRPLBGSD 240
 Db 181 LLLIQLVLSCTSDRSPLPSETIHDNCPSSASFLSRITFWWTGILVRGYRPLBGSD 240
 QY 241 LMSLNKEDTSBOVPVLYKMKKKCACTRRQOPVNVVSSQDPAQPKSSKYDAEEVEAL 300
 Db 241 LMSLNKEDTSBOVPVLYKMKKKCACTRRQOPVNVVSSQDPAQPKSSKYDAEEVEAL 300
 QY 301 IVKSPQKKNPSPLPKVLKTFGPFYFLMSFFFKAIHDLMEFSGPOLIKLILKFVNDTAPD 360
 Db 301 IVKSPQKKNPSPLPKVLKTFGPFYFLMSFFFKAIHDLMEFSGPOLIKLILKFVNDTAPD 360
 QY 361 WQGYFTVLLFVTAQCTLVLAHQYFHLCPVSGMKRIKTAIVGAVYRKALVITNSARKSTV 420
 Db 361 WQGYFTVLLFVTAQCTLVLAHQYFHLCPVSGMKRIKTAIVGAVYRKALVITNSARKSTV 420
 QY 421 GEIVNLSVDAQRMDLATYINMWSAPLVYLLALYLLMLNLSVLAQVAVVLMVFN 480
 Db 421 GEIVNLSVDAQRMDLATYINMWSAPLVYLLALYLLMLNLSVLAQVAVVLMVFN 480
 QY 481 AVNMAKTKTYQVAHMSKDNRIKLMNEILNGIKVLYAMELAFKDKVLAIROBELKYLK 540
 Db 481 AVNMAKTKTYQVAHMSKDNRIKLMNEILNGIKVLYAMELAFKDKVLAIROBELKYLK 540
 QY 541 KSAVLASAVGTFTWCTPFLVALCTFAVYVTIDENNILDAQTAFAVSLAFNLRPPLNLP 600
 Db 541 KSAVLASAVGTFTWCTPFLVALCTFAVYVTIDENNILDAQTAFAVSLAFNLRPPLNLP 600
 QY 601 MVTSSIVQASVLSKRLRIFLSHELEPDSIERRPVQKGGGNTSITVENAFTMARSDPT 660
 Db 601 MVTSSIVQASVLSKRLRIFLSHELEPDSIERRPVQKGGGNTSITVENAFTMARSDPT 660
 QY 661 LMGITFSPGALVAVVGQVCGGKSLSLALLMEQDVEGHVMIKGSVAVYPOQAMQND 720
 Db 661 LMGITFSPGALVAVVGQVCGGKSLSLALLMEQDVEGHVMIKGSVAVYPOQAMQND 720
 QY 721 SLRENILFGQLEBPYRYSVIOACALLPDLIELPSGDRTEIGEGVNLSSGQKQVSLAR 780
 Db 721 SLRENILFGQLEBPYRYSVIOACALLPDLIELPSGDRTEIGEGVNLSSGQKQVSLAR 780
 QY 781 AVYSNADTYLPDPLSAVDNAHVGKHIFENYIGRGMKAKTRILVTHSMSTLPQVDYIIV 840
 Db 781 AVYSNADTYLPDPLSAVDNAHVGKHIFENYIGRGMKAKTRILVTHSMSTLPQVDYIIV 840
 QY 841 MSGGKTSEMGSYOELLARDGAFAPFLRTVASTQEOBAEENGTVGSGPKKAKOMENGM 900
 Db 841 MSGGKTSEMGSYOELLARDGAFAPFLRTVASTQEOBAEENGTVGSGPKKAKOMENGM 900
 QY 901 LVYDSAGKQLOROLSSSSSYSGDISRHNNSTAELOKAEKKEBTWKLMEADKAKQGVKL 960
 Db 901 LVYDSAGKQLOROLSSSSSYSGDISRHNNSTAELOKAEKKEBTWKLMEADKAKQGVKL 960
 QY 961 SVYWDYKAKAIGLIFSLSTIFLFCNHNVSALASNYWLSLWTDPIVNGTQSHYVRLSVYG 1020
 Db 961 SVYWDYKAKAIGLIFSLSTIFLFCNHNVSALASNYWLSLWTDPIVNGTQSHYVRLSVYG 1020
 QY 1021 ALGISGIAVFGYSMAVSTGIIASRCIHDVLDHSLIRSMSPFERPSCNVLNRSKEL 1080
 Db 1021 ALGISGIAVFGYSMAVSTGIIASRCIHDVLDHSLIRSMSPFERPSCNVLNRSKEL 1080
 QY 1081 DTVDSMIPETVYKMFMSLFNVIGACIVILATPIAIIIPPLGLIYFFVORFYVASSROL 1140
 Db 1081 DTVDSMIPETVYKMFMSLFNVIGACIVILATPIAIIIPPLGLIYFFVORFYVASSROL 1140
 QY 1141 KRLSVSRSPYSHFNMTLLIGSVITAPFEQERFIHOSDLKVDENOKAAYPSIVANRWLA 1200
 Db 1141 KRLSVSRSPYSHFNMTLLIGSVITAPFEQERFIHOSDLKVDENOKAAYPSIVANRWLA 1200
 QY 1201 VRLCYGNCIVLPAALPAVYSRSLSGVLGSLVSLOVTTVLMNLVRRSSMETNIVA 1260

Db 1201 VRLCEVGNCLVLPALFAVVISRHSLSAGLVGLSVSYSLQVTTYLTNMLVMSSEMETNIVA 1260
Qy 1261 VERLKEVSETEKEAPWQIQETAPSPSSPQVGRVEFRNRYCLRYREDLDVFLHINVTINGG 1320
Db 1261 VERLKEVSETEKEAPWQIQETAPSPSSPQVGRVEFRNRYCLRYREDLDVFLHINVTINGG 1320
Qy 1321 ERVGVIGRTGAGKSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
Db 1321 ERVGVIGRTGAGKSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
Qy 1381 SCSLRMNDPFSQVSDSEEWMTSLAHKIDFVSAIPDLHDHRCAGGNNLSVGRQIVCL 1440
Db 1381 SCSLRMNDPFSQVSDSEEWMTSLAHKIDFVSAIPDLHDHRCAGGNNLSVGRQIVCL 1440
Qy 1441 ARALLRKTIIVLDEATAVAVLETDLLIQSTIRIQFEDCTVLTIAHRLNTIMDYRVIVL 1500
Db 1441 ARALLRKTIIVLDEATAVAVLETDLLIQSTIRIQFEDCTVLTIAHRLNTIMDYRVIVL 1500
Qy 1501 DKGEIQEYGAPSDLLIQRGFLFSMAKDAGLV 1531
Db 1501 DKGEIQEYGAPSDLLIQRGFLFSMAKDAGLV 1531

RESULT 13

ADB87954 standard; protein; 1531 AA.

ADB87954;

04-DEC-2003 (first entry)

Human UGT1A1 protein sequence SEQ ID NO:678.

irrinotecan; cancer; UGT1A1; cytosolic; topoisomerase I inhibitor;
colorectal cancer; cervical cancer; gastric cancer; lung cancer;
ovarian cancer; pancreatic cancer; malignant glioma;
uridine diphosphate glycosyltransferase1 member A1.

Homo sapiens.

WO2003013536-A2.

20-FEB-2003.

23-JUL-2002; 2002WO-EP008217.

23-JUL-2001; 2001EP-00117608.

24-MAY-2002; 2002EP-00011710.

(EPID-) EPIDAUROS BIOTECHNOLOGIE AG.

Heinrich G, Kerb R;

WPI; 2003-289896/28.

Use of irinotecan to treat cancer patient by determining if patient has
variant alleles of UGT1A1 gene, administering increased/decreased amounts
of irinotecan based on increased/decreased levels of UGT1A1 gene product.

Disclosure: SEQ ID NO 678; 107bp; English.

The invention relates to the novel use of irinotecan to treat a patient
suffering from cancer. This involves determining if the patient has one
or more variant alleles of the UGT1A1 gene, and if the patient has one
or more of such variant alleles, irinotecan is administered in an increased
or decreased amount in comparison to the amount that is administered
without regard to the patient's alleles in the UGT1A1 gene. The invention
has cytosolic activity. A composition of the invention acts as a
topoisomerase I inhibitor. The method is useful for treating a patient,
an animal e.g. mouse or a human, preferably African or Asian, suffering
from cancer such as colorectal, cervical, gastric cancer, lung, ovarian,
pancreatic cancer or malignant glioma. The present sequence is used in
the exemplification of the invention.

XX Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 7; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRGCSADGSDPDLMDMNTWNTSNPDFTKCFONTYLVWVPCFYLAACPFYLYLSRH 60
Db 1 MALRGCSADGSDPDLMDMNTWNTSNPDFTKCFONTYLVWVPCFYLAACPFYLYLSRH 60
Qy 61 DRGYQMTPLNKKTALGFLIMYICWADLFPSFMRSGRGLFLAVFLVSPYLLGITTLLA 120
Db 61 DRGYQMTPLNKKTALGFLIMYICWADLFPSFMRSGRGLFLAVFLVSPYLLGITTLLA 120
Qy 121 TFLIQLEBRKGVSSGIMLTFEVLVAVCALALISKIMTALKEDAOVDLFFDITFYVYFS 180
Db 121 TFLIQLEBRKGVSSGIMLTFEVLVAVCALALISKIMTALKEDAOVDLFFDITFYVYFS 180
Qy 181 LLLIQVLVSCFSDRSPLESETIHDNPPCESSASFLSRITFWMITGLIVRGYROPLEGSD 240
Db 181 LLLIQVLVSCFSDRSPLESETIHDNPPCESSASFLSRITFWMITGLIVRGYROPLEGSD 240
Qy 241 LMSLNKEDTSFOVVPVLYKMKKECAKTRKQPVKVVYSSKQPAQKSSKYDANEVVAL 300
Db 241 LMSLNKEDTSFOVVPVLYKMKKECAKTRKQPVKVVYSSKQPAQKSSKYDANEVVAL 300
Qy 301 IVKSPQKEMNPSLKFVLYKTFGPFYFMSFFPKALHDLMMFSGPOLKLLIKFVNDTKAPD 360
Db 301 IVKSPQKEMNPSLKFVLYKTFGPFYFMSFFPKALHDLMMFSGPOLKLLIKFVNDTKAPD 360
Qy 361 WQGFYFVLLFVTAQCLQTLVHQYFHI CFVSGMKIKTAVIGAVYRKALVITNSARKSSTV 420
Db 361 WQGFYFVLLFVTAQCLQTLVHQYFHI CFVSGMKIKTAVIGAVYRKALVITNSARKSSTV 420
Qy 421 GEIVNLSVDAQRPMDLATYINMTWSAPLOVITLXILMLGSGVLAVGVWMLMVVN 480
Db 421 GEIVNLSVDAQRPMDLATYINMTWSAPLOVITLXILMLGSGVLAVGVWMLMVVN 480
Qy 481 AVMMKRTYOVAAHMKSKDNRIKIMNELINGIKVLYKYLAMELAFKDYALIROBELKYLK 540
Db 481 AVMMKRTYOVAAHMKSKDNRIKIMNELINGIKVLYKYLAMELAFKDYALIROBELKYLK 540
Qy 541 KSAVLSAVGTFTWCTPFLVALCTPAVYVTIDENNIIIDAQAFVSLAFNLRPLNLTLP 600
Db 541 KSAVLSAVGTFTWCTPFLVALCTPAVYVTIDENNIIIDAQAFVSLAFNLRPLNLTLP 600
Qy 601 MYSSIVASVSLKRLRIFLSHELEBPSIRRRPVKGGGINSITVRATFTWARSDDPT 660
Db 601 MYSSIVASVSLKRLRIFLSHELEBPSIRRRPVKGGGINSITVRATFTWARSDDPT 660
Qy 661 LMGTFPSIPEGALVAVVGVQVCGKSSLLSALLAEMDKVEGHVALIKGSAVYVPOQAMIQND 720
Db 661 LMGTFPSIPEGALVAVVGVQVCGKSSLLSALLAEMDKVEGHVALIKGSAVYVPOQAMIQND 720
Qy 721 SLRENIILFCQLEBRYRSVIOACALLPDLILPSGDRTEIGKGVNLSGGOKRVSLAR 780
Db 721 SLRENIILFCQLEBRYRSVIOACALLPDLILPSGDRTEIGKGVNLSGGOKRVSLAR 780
Qy 781 AVYSNADILYLPDDPLSAVDAAVKGKIFENYVIGPKMLKNKRIILVTHMSVLPQVDVIV 840
Db 781 AVYSNADILYLPDDPLSAVDAAVKGKIFENYVIGPKMLKNKRIILVTHMSVLPQVDVIV 840
Qy 841 MSGKISMGVSYOELIANDGAFAEFLRTYASTBOBDAEENGVTGVSPGKEAKOMENGM 900
Db 841 MSGKISMGVSYOELIANDGAFAEFLRTYASTBOBDAEENGVTGVSPGKEAKOMENGM 900
Qy 901 LVTSAGKOLROULSSSSYSGDISRRHNSYAELOKAKKEETWKLMEADKAGQGYVL 960
Db 901 LVTSAGKOLROULSSSSYSGDISRRHNSYAELOKAKKEETWKLMEADKAGQGYVL 960
Qy 961 SYVWDYKALIGLFLISFLSIFPMCHVSGALASNYVLSLWTDPIYNGQERTKVLSTYVG 1020
Db 961 SYVWDYKALIGLFLISFLSIFPMCHVSGALASNYVLSLWTDPIYNGQERTKVLSTYVG 1020

Db 961 SVYWDYKAIAGLFIISFISIFLMCNHVSALASNYWLSIMTDDPIVNGIOETKVRLSVYG 1020
QY 1021 ALGISOGIAVGYMAVSIIGIIASRCHVDLHSIIASRSPSPFERPPSGULVNRFSKEL 1080
Db 1021 ALGISOGIAVGYMAVSIIGIIASRCHVDLHSIIASRSPSPFERPPSGULVNRFSKEL 1080
QY 1081 DTVDMSIPEVIAKMGMSLFVNIIGACIYLLATPIAAIIIPPLGIIYFVQGFYVASSROL 1140
Db 1081 DTVDMSIPEVIAKMGMSLFVNIIGACIYLLATPIAAIIIPPLGIIYFVQGFYVASSROL 1140
QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANWMLA 1200
Db 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANWMLA 1200
QY 1201 VRLECVNCTIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYTNMVLVMSSEMETNIVA 1260
Db 1201 VRLECVNCTIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYTNMVLVMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMOIOETAPSPSWPQVGRVFRNYCLRYREDLPVLJRHINVTINGG 1320
Db 1261 VERLKEYSETEKEAPMOIOETAPSPSWPQVGRVFRNYCLRYREDLPVLJRHINVTINGG 1320
QY 1321 EKVGIIVGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGHDLRFKTIIPQDPVLF 1380
Db 1321 EKVGIIVGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGHDLRFKTIIPQDPVLF 1380
QY 1381 SSGSRMNLDPFSQYSDSEVWTSLELAHLKDFVSLPDKLDHECAGEGNSVGOQOLVCL 1440
Db 1381 SSGSRMNLDPFSQYSDSEVWTSLELAHLKDFVSLPDKLDHECAGEGNSVGOQOLVCL 1440
QY 1441 ABALLRKTIXLVDEATAVADLETDDLIQSTIRTOFEDCTVLTARHLANTIMDTRYIVL 1500
Db 1441 ABALLRKTIXLVDEATAVADLETDDLIQSTIRTOFEDCTVLTARHLANTIMDTRYIVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMAXDAGLV 1531
Db 1501 DKGEIOEYGAPSDLLQORGLFYSMAXDAGLV 1531

RESULT 14
ADB96937 standard; protein; 1531 AA.
XX ADB96937;
AC ADB96937;
DT 04-DEC-2003 (first entry)
XX
DE Human MDR1 related protein sequence SEQ ID NO:678.
XX
KW Irinotecan; colorectal cancer; cervical cancer; gastric cancer;
KW lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
KW multidrug resistance 1; MDR1; cytosolic; human; Cyp3A5; MRP1; MDR1;
KW TDP1.
XX
OS Homo sapiens.
XX
PN MO2003013537-A2.
XX
PD 20-FEB-2003.
XX
PF 23-JUL-2002; 2002MO-EP008218.
XX
PR 23-JUL-2001; 2001EP-00117608.
XX
PR 24-MAY-2002; 2002EP-00011710.
XX
PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI Heinrich G, Kerb R;
XX
DR WPI; 2003-268145/26.
XX
PT New use of irinotecan for preparation of pharmaceutical compositions for
PT treating cancer in subject having genome with variant allele comprising

PT multidrug resistance 1 polynucleotide.
XX
PS Disclosure; SEQ ID NO 678; 130pp; English.
XX
CC The invention relates to the novel use of irinotecan or its derivative
CC for the preparation of pharmaceutical compositions for treating
CC colorectal, cervical, gastric, lung, ovarian or pancreatic cancer, or
CC malignant glioma in a subject having a genome with a variant allele which
CC comprises a multidrug resistance 1 (MDR1) polynucleotide. A composition
CC of the invention has cytostatic activity. The invention is useful for the
CC preparation of pharmaceutical compositions for treating colorectal,
CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant
CC glioma in a subject (preferably human, more preferably African or Asian)
CC or a mouse. The present sequence is used in the exemplification of the
CC invention.
XX
SQ Sequence 1531 AA;
XX
Query Match 79.4%; Score 7860; DB 7; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSADGSDPLMDMNTWNTSNPDFTKCFQNTLVVWPCFYLMACEPFYLYLSRH 60
Db 1 MALRGFCSADGSDPLMDMNTWNTSNPDFTKCFQNTLVVWPCFYLMACEPFYLYLSRH 60
QY 61 DRGVIQMTPLNKTXTALGFLIMIVCWADLFYSFWERSRGIFLAPVLVSPTLIGITTLA 120
Db 61 DRGVIQMTPLNKTXTALGFLIMIVCWADLFYSFWERSRGIFLAPVLVSPTLIGITTLA 120
QY 121 TPLIOLERRRGVSSGIMLTFWLVALVCALAIIRSKTMTLXKDAQVLDPRDITFYVFS 180
Db 121 TPLIOLERRRGVSSGIMLTFWLVALVCALAIIRSKTMTLXKDAQVLDPRDITFYVFS 180
QY 181 LLLIQLVLSGFSRSPPLFSEITIHDPNCPSSASFLSRITFWMTGLIVRGYRPLEGSD 240
Db 181 LLLIQLVLSGFSRSPPLFSEITIHDPNCPSSASFLSRITFWMTGLIVRGYRPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPQPKSSKVDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPQPKSSKVDANEVEAL 300
QY 301 IVKSPQKMNPSLFKYLKTFGYPFLMSFFPKAIDHLMSSGQILLLKFPVNDTAPD 360
Db 301 IVKSPQKMNPSLFKYLKTFGYPFLMSFFPKAIDHLMSSGQILLLKFPVNDTAPD 360
QY 361 WQGYFTVLLFVTAQTLVLAHQYFHLCFVSGMRKTAIVGAYRKALVITNSARKSSTV 420
Db 361 WQGYFTVLLFVTAQTLVLAHQYFHLCFVSGMRKTAIVGAYRKALVITNSARKSSTV 420
QY 421 GEIVNIMSVDQSPMDLAIYINMWSAPLOVILALYLLMLNLPISVLAVAVWLVVNV 480
Db 421 GEIVNIMSVDQSPMDLAIYINMWSAPLOVILALYLLMLNLPISVLAVAVWLVVNV 480
QY 481 AVNAAKTKTYQVAHMSKONRITLNMHILNGIVLKLXAMELAFKQKVLAIROBELKVLK 540
Db 481 AVNAAKTKTYQVAHMSKONRITLNMHILNGIVLKLXAMELAFKQKVLAIROBELKVLK 540
QY 541 KSAVLSAVGFTFWCTPFLVALCTPAVYVYTIIDENNLDAQTAFVSLAFNULRPLNLI 600
Db 541 KSAVLSAVGFTFWCTPFLVALCTPAVYVYTIIDENNLDAQTAFVSLAFNULRPLNLI 600
QY 601 MVISSIVQASVSLKRLRIFLSHELEPDSIERRPVDDGGTNSITVNNATFTYARSDPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHELEPDSIERRPVDDGGTNSITVNNATFTYARSDPT 660
QY 661 LMGITFSIPGALVAVVGVGCGKSSLSLALXEMDKVEGHVAIKGSVAVYPOQAMQND 720
Db 661 LMGITFSIPGALVAVVGVGCGKSSLSLALXEMDKVEGHVAIKGSVAVYPOQAMQND 720
QY 721 SURENLLFGQLEPEYRSVYIQAALLPDLIELPSDREITIGKGVNLSCGQORVSLAR 780
Db 721 SURENLLFGQLEPEYRSVYIQAALLPDLIELPSDREITIGKGVNLSCGQORVSLAR 780

QY 781 AVYSNADIIYLFDDPLSAVDHVGKHFENYIGPKMLKNKTRIIIVTSHMSYLPQVDVII 840
DB 781 AVYSNADIIYLFDDPLSAVDHVGKHFENYIGPKMLKNKTRIIIVTSHMSYLPQVDVII 840
QY 841 MSGGKISEMGSYOBELARDGAFAEFLRTYASTBODSENGVGVSPGKEAKOMEGM 900
DB 841 MSGGKISEMGSYOBELARDGAFAEFLRTYASTBODSENGVGVSPGKEAKOMEGM 900
QY 901 LVTDAGKOLOROLSSSSSSSGDISRHHNSTAELOKAKKEETKLEADKAOCTQYKL 960
DB 901 LVTDAGKOLOROLSSSSSSSGDISRHHNSTAELOKAKKEETKLEADKAOCTQYKL 960
QY 961 SVYDYMKAIGLFTISLIFLMCNHVSALASNYWLSMTDDPIVNGTOEHTKVALSYG 1020
DB 961 SVYDYMKAIGLFTISLIFLMCNHVSALASNYWLSMTDDPIVNGTOEHTKVALSYG 1020
QY 1021 ALGISOGIAVRGYSMAVSTIGIILASRCHVDLHSHILSPSPFERTSSGMLVNRFSKL 1080
DB 1021 ALGISOGIAVRGYSMAVSTIGIILASRCHVDLHSHILSPSPFERTSSGMLVNRFSKL 1080
QY 1081 DTVDSMIPEVIKMFMSLFENVIGACIVILLATPIAIIIPPLGLIYFVQRFYVASSRQL 1140
DB 1081 DTVDSMIPEVIKMFMSLFENVIGACIVILLATPIAIIIPPLGLIYFVQRFYVASSRQL 1140
QY 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENQKAYYSIVANRWLA 1200
DB 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENQKAYYSIVANRWLA 1200
QY 1201 VRLSCVNCICVLPALFAVIRSRHSLASGLVSGVSYLOVTTYLMVLVRMSEMETNIVA 1260
DB 1201 VRLSCVNCICVLPALFAVIRSRHSLASGLVSGVSYLOVTTYLMVLVRMSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPPSMPQGVVEFRNYCLARYEDLDEVLRHINVTINGG 1320
DB 1261 VERLKEYSETEKEAPWQIOETAPPSMPQGVVEFRNYCLARYEDLDEVLRHINVTINGG 1320
QY 1321 EKVGI VGTGAGKSSLTGLFRINSAEGEIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
DB 1321 EKVGI VGTGAGKSSLTGLFRINSAEGEIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
QY 1381 SGSLRMLNDPFSQYSDSEVMTSLAHLKDPYSALPDLDHHCAGGENLSVGORQVLC 1440
DB 1381 SGSLRMLNDPFSQYSDSEVMTSLAHLKDPYSALPDLDHHCAGGENLSVGORQVLC 1440
QY 1441 ARALLRKTIIIVLEATAVADLETDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYRYVIL 1500
DB 1441 ARALLRKTIIIVLEATAVADLETDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYRYVIL 1500
QY 1501 DKGEIOEYGAPSDLLIQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGAPSDLLIQORGLFYSMAKDAGLV 1531
RESULT 15
ADB92128
ID ADB92128 standard; protein; 1531 AA.
XX
AC ADB92128;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human MDRI related protein sequence SEQ ID NO:678.
XX
KW irinotecan; colorectal cancer; cervical cancer; gastric cancer;
KW lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
KW multidrug resistance 1; MDRI; cytosolic; human; UGT1A1; MRP1; TOP1.
XX
OS Homo sapiens.
XX
PN MO2003013535-A2.
XX
PD 20-FEB-2003.

XX 23-JUN-2002; 2002MO-EP008220.
PF 23-JUN-2001; 2001EP-00117608.
XX 24-MAY-2002; 2002EP-00011710.
PR (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX Heinrich G, Korb R;
XX WPI; 2003-342400/32.
DR
XX
XX New use of irinotecan for preparation of pharmaceutical compositions for
PT treating cancer in subject having genome with variant allele comprising
PT multidrug resistance 1 polynucleotide.
XX
XX Disclosure; SEQ ID NO 678; 104pp; English.
PS
XX
XX The invention relates to a novel use of irinotecan or its derivative for
CC the preparation of a pharmaceutical composition for treating colorectal,
CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant
CC glioma in a subject having a genome with a variant allele which comprises
CC a multidrug resistance 1 (MDRI) polynucleotide. A composition of the
CC invention has cytostatic activity. The present sequence is used in the
CC exemplification of the invention.
XX
XX Sequence 1531 AA;
SQ
Query Match 79.4%; Score 7860; DB 7; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFSGADSGDSDPLMDMNTWNTSNPDFTKCFQNTVYLWVPCFYLMACPPFYLYLSRH 60
DB 1 MALRGFSGADSGDSDPLMDMNTWNTSNPDFTKCFQNTVYLWVPCFYLMACPPFYLYLSRH 60
QY 61 DRGIIQMTPLNKTATGFLIMIVCMADLFYSFMRSGRGTFLAVPLVPLTGITTLA 120
DB 61 DRGIIQMTPLNKTATGFLIMIVCMADLFYSFMRSGRGTFLAVPLVPLTGITTLA 120
QY 121 TFLIQLERRKGVQSSGIMLTFWLVALCALILSKIMTALKEBQVDFPDITFYVFS 180
DB 121 TFLIQLERRKGVQSSGIMLTFWLVALCALILSKIMTALKEBQVDFPDITFYVFS 180
QY 181 LLLIQLVLSCFSDSPFSETIHDNPPCESASFLSRITFWITGLIVRGYROPLESGD 240
DB 181 LLLIQLVLSCFSDSPFSETIHDNPPCESASFLSRITFWITGLIVRGYROPLESGD 240
QY 241 LMSLNKEDTSEQVVPVLYKMKKSCAKTRKQPVKVYVSSKDPAPKESKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLYKMKKSCAKTRKQPVKVYVSSKDPAPKESKVDANEVEAL 300
QY 301 IVKSPQKMNPSLFRVLYKTGFPYFMSFPFKAIHDLMPFSGPOTIKLLIFVNDTKAPD 360
DB 301 IVKSPQKMNPSLFRVLYKTGFPYFMSFPFKAIHDLMPFSGPOTIKLLIFVNDTKAPD 360
QY 361 WQGYFYTVLFTVTAQLQTLVLHQYFHLCEVSGMBIKTAVIGAVYRKALVITNSARKSSTV 420
DB 361 WQGYFYTVLFTVTAQLQTLVLHQYFHLCEVSGMBIKTAVIGAVYRKALVITNSARKSSTV 420
QY 421 GEIYVNLMSVDAQRFMDLATYINMTWSAPLQVITLALYLLMLMGSVLAVGVAVMLVPEVN 480
DB 421 GEIYVNLMSVDAQRFMDLATYINMTWSAPLQVITLALYLLMLMGSVLAVGVAVMLVPEVN 480
QY 481 AVMMKTKTYGVAAHKSNDNR IKLMNELNLNGIYVLYKLYAMELAFDXYLAROBELXYLK 540
DB 481 AVMMKTKTYGVAAHKSNDNR IKLMNELNLNGIYVLYKLYAMELAFDXYLAROBELXYLK 540
QY 541 KSAVLSAVGFTFWCTPFLVALCTFAVYVITDENNIIIDAQTAFYSLAFNLIRPLINTIP 600
DB 541 KSAVLSAVGFTFWCTPFLVALCTFAVYVITDENNIIIDAQTAFYSLAFNLIRPLINTIP 600
QY 601 MVISSIVQSVSLKRLRIFLSHELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660

Db 601 MVISSIVQASVSLKRLRIFLSHBELEPDSIRRRPVKQDGTNSITVRNATFTWARSDEPT 660
 QY 661 LNCITFSPREGALVAVVQVCCGKSSLSALLAEMDKYGHVAIKGSVAVYPOQAMIOND 720
 Db 661 LNCITFSPREGALVAVVQVCCGKSSLSALLAEMDKYGHVAIKGSVAVYPOQAMIOND 720
 QY 721 SLRENIIFGCOLLEBPYRSVIQACALPDLFILPSGDRTEIGBKGVNLSSGQOKORVSLAR 780
 Db 721 SLRENIIFGCOLLEBPYRSVIQACALPDLFILPSGDRTEIGBKGVNLSSGQOKORVSLAR 780
 QY 781 AVYNSADIYLPDDPLSAVDHVGKRIPEENVIGPKMLNKXTRILVTHSMXYLPQVDYIIV 840
 Db 781 AVYNSADIYLPDDPLSAVDHVGKRIPEENVIGPKMLNKXTRILVTHSMXYLPQVDYIIV 840
 QY 841 MSGKISMSGSYQELIADGAPAEFLRTYASTEOQDAENGVTGSGPGKEAKOMENGM 900
 Db 841 MSGKISMSGSYQELIADGAPAEFLRTYASTEOQDAENGVTGSGPGKEAKOMENGM 900
 QY 901 LVTSAGKQLOROSLSSSSSGSDISRHHNSTAELOKAKKEETMKLEADKAOQGVKL 960
 Db 901 LVTSAGKQLOROSLSSSSSGSDISRHHNSTAELOKAKKEETMKLEADKAOQGVKL 960
 QY 961 SVYWDYMKALGLFISFLSIFLMCNHVSALASNYWLSLMTDDPIVNGTOEHTKYRLSYG 1020
 Db 961 SVYWDYMKALGLFISFLSIFLMCNHVSALASNYWLSLMTDDPIVNGTOEHTKYRLSYG 1020
 QY 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLHSLRSFMSFFERTPSGNLVNRFSEKL 1080
 Db 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLHSLRSFMSFFERTPSGNLVNRFSEKL 1080
 QY 1081 DTVUSMTPEVIKMMGSLFENVIGACIVILLATPIAIIIPPLGIYFFVQRFYVASSROL 1140
 Db 1081 DTVUSMTPEVIKMMGSLFENVIGACIVILLATPIAIIIPPLGIYFFVQRFYVASSROL 1140
 QY 1141 KRLESVRSPIYSHFNETILGVSIVIRAFEOERFIHOSDLKYDENOKAYYPSIVANRWLA 1200
 Db 1141 KRLESVRSPIYSHFNETILGVSIVIRAFEOERFIHOSDLKYDENOKAYYPSIVANRWLA 1200
 QY 1201 VRELCVGNCLVLPALPAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRRMSEMETNIVA 1260
 Db 1201 VRELCVGNCLVLPALPAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRRMSEMETNIVA 1260
 QY 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVFRNVCILAYREDLDPVLRIHINVTINGG 1320
 Db 1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVFRNVCILAYREDLDPVLRIHINVTINGG 1320
 QY 1321 EKVGIIVERTGAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRPKTIIPQDPVLF 1380
 Db 1321 EKVGIIVERTGAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRPKTIIPQDPVLF 1380
 QY 1381 SGLRMLNDPPSOYSDEBEWTSLELAHLKDFVSALPDKLDHECAEGGENLSVGQROLVCL 1440
 Db 1381 SGLRMLNDPPSOYSDEBEWTSLELAHLKDFVSALPDKLDHECAEGGENLSVGQROLVCL 1440
 QY 1441 ARALLRKTILVDEBATAVLETDLLOSTIRTOFEDCTVLTARLNTIMDYTRIVYL 1500
 Db 1441 ARALLRKTILVDEBATAVLETDLLOSTIRTOFEDCTVLTARLNTIMDYTRIVYL 1500
 QY 1501 DKGEIOEXGAPSDLLQORGLFYMAKDAGLV 1531
 Db 1501 DKGEIOEXGAPSDLLQORGLFYMAKDAGLV 1531

Search completed: December 15, 2005, 15:17:03
 Job time : 186.366 secs

THE UNIVERSITY OF CHICAGO

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:57:04 ; Search time 37.8285 Seconds
(without alignments)
4901.320 Million cell updates/sec

Title: US-10-665-283-1

Perfect score: 9903
Sequence: 1 MALRGFCSADGSDPLMDMNV.....RSVAVAKKPKFSIPDLS 1927

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*

1: p1r1.*
2: p1r2.*
3: p1r3.*
4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7849	79.3	1531	1	DVHUR
2	4464.5	45.3	1527	2	JB036
3	3587	36.2	1545	1	S71841
4	3507.5	35.4	1541	1	S71839
5	3289	33.2	1494	2	B89447
6	3225	32.6	1573	2	T21219
7	3195.5	32.3	1502	2	T42216
8	2838	28.7	1515	1	S51863
9	2732	27.6	1398	2	T20434
10	2548	25.7	1478	2	T38712
11	2384.5	24.1	1623	2	T01369
12	2356.5	23.8	1622	2	D86428
13	2326	23.5	1495	2	B86428
14	2294.5	23.3	1144	2	T27408
15	2263.5	22.9	1559	1	S64757
16	2263.5	22.9	1488	2	B86428
17	2238.5	22.6	1516	2	F84919
18	2230	22.5	1539	2	T48059
19	2144.5	21.7	1335	2	T00961
20	2144.5	21.7	1514	2	T52080
21	2140	21.6	1515	2	T52081
22	2131.5	21.5	1490	2	T47840
23	2129	21.5	1545	2	T46645
24	2123	21.4	1153	2	T26883
25	2109	21.3	1545	2	T42751
26	2104.5	21.3	1511	2	T42711
27	2098	21.2	1546	2	T42728
28	2069.5	20.9	1389	2	T47796
29	2062.5	20.8	1661	2	S64800

30	2037	20.6	1121	2	C87973	protein Y43F8C.12
31	2014.5	20.3	1582	2	A56248	sulfonylurea recep
32	2008	20.3	1592	2	S48933	probable transpor
33	2007	20.3	390	2	S68403	inward rectifier p
34	1992	20.1	390	2	UC4689	inwardly rectifyin
35	1935	19.5	390	2	A57616	inward rectifier K
36	1932	19.5	390	2	UC7901	inwardly-rectifyin
37	1888	19.1	1477	2	S64616	YORI protein - Yea
38	1884.5	19.0	1037	2	T50518	ABC transporter-1i
39	1805	18.2	1146	2	F84487	probable ABC trans
40	1722	17.4	946	1	UC5667	multidrug resistan
41	1693	17.1	1548	1	DVLNS	multidrug resistan
42	1687	17.0	1421	2	T34225	hypothetical prote
43	1606	16.2	1427	2	T20903	hypothetical prote
44	1575.5	15.9	1427	2	T39219	atp-binding caset
45	1558	15.7	1469	2	T50210	probable ABC trans

ALIGNMENTS

RESULT 1

DVHUR
multidrug resistance protein (cell line H69AR) - human
N:Alternate names: multidrug resistance-associate protein (MRP)
C:Species: Homo sapiens (man)
C>Date: 30-Jun-1993 #sequence_revision 05-Dec-1998 #text_change 19-Jan-2001
C/Accession: A44231; A37495
R/Cole, S.P.C.; Bhaidwaj, G.; Gerlach, J.H.; Mackie, J.E.; Grant, C.E.; Almquist, K.C.,
Science 258, 1650-1654, 1992
A/Title: Overexpression of a transporter gene in a multidrug-resistant human lung cancer
A/Reference number: A44231, MUID:9308080, PMID:1360704
A/Accession: A44231
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1531
A/Cross-references: UNIPARC:UPI00001746C; GB:105628; NID:91835658
A/Experimental source: small cell lung carcinoma cell line H69AR
A/Note: Sequence extracted from NCBI backbone (NCBI:119851); this sequence has been co
R/Cole, S.P.C.; Deele, R.G.
Science 260, 879, 1993
A/Title: Multidrug resistance-associate protein: sequence correction.
A/Reference number: A37495; MUID:93262415; PMID:8098549
A/Accession: A37495
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-60 <CO2>
A/Cross-references: UNIPARC:UPI00001746C; GB:105628; NID:91835658
A/Note: sequence extracted from NCBI backbone (NCBI:131929)
C/Genetic:
A/Gene: GDB:MRP
A/Cross-references: GDB:136335; OMIM:158343
A/Map position: 16p13.1-16p13.1
C/Map family: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C/Keywords: antibiotic resistance; ATP; duplication; nucleotide binding; P-loop; transmem
F/61-84/Domain: ATP-binding cassette homology <ABC1>
F/67-68/Region: nucleotide-binding motif A (P-loop)
F/78-79/Region: nucleotide-binding motif B
F/110-150/Domain: ATP-binding cassette homology <ABC2>
F/127-133/Region: nucleotide-binding motif A (P-loop)
F/150-145/Region: nucleotide-binding motif B

Query Match 79.3%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	MALRGFCSADGSDPLMDMNVNTNTSNPDFFKCRONTLVWVPCYVLMACPFYFLYSRH	60
Db	1	MALRGFCSADGSDPLMDMNVNTNTSNPDFFKCRONTLVWVPCYVLMACPFYFLYSRH	60
Qy	61	DRGVQMTPLNKTALGFLMTVCVADLFYFSPERSRGIFLAPVPLVSPFLTGITTLA	120
Db	61	DRGVQMTPLNKTALGFLMTVCVADLFYFSPERSRGIFLAPVPLVSPFLTGITTLA	120

```

QY 121 TFLIQERRKGVSSGIMLTFLMVALYCALAIIIRSKIMTALKEDAQVDLFEDITFYVYS 180
D 121 TFLIQERRKGVSSGIMLTFLMVALYCALAIIIRSKIMTALKEDAQVDLFEDITFYVYS 180
QY 181 LLLIQVLVISCSDSPFLFSETIHDNPPCBSSASFLSIITFWITGLIVRGYRQPLESD 240
D 181 LLLIQVLVISCSDSPFLFSETIHDNPPCBSSASFLSIITFWITGLIVRGYRQPLESD 240
QY 241 LMSINKEDTSBOVVPVLVKNMKKECAKTRKOPVKVYVSSKDPAPCKESSKYDANEVVAL 300
D 241 LMSINKEDTSBOVVPVLVKNMKKECAKTRKOPVKVYVSSKDPAPCKESSKYDANEVVAL 300
QY 301 IVKSPKEMNPSLFKVLVYKTFGPYFLMSFFPKAIIHDLMMFSGPOLIKLIRFVNDTKAPD 360
D 301 IVKSPKEMNPSLFKVLVYKTFGPYFLMSFFPKAIIHDLMMFSGPOLIKLIRFVNDTKAPD 360
QY 361 WQGFYTYVLLFVTACLQTLVHOYFHICFVSGMRKTKVIGAVRKALVITNSARKSSTV 420
D 361 WQGFYTYVLLFVTACLQTLVHOYFHICFVSGMRKTKVIGAVRKALVITNSARKSSTV 420
QY 421 GEIYNLMSVDQRFPMDLATYINMISAPLOVILALYLMNLGSPVLGAVVMVLMVFN 480
D 421 GEIYNLMSVDQRFPMDLATYINMISAPLOVILALYLMNLGSPVLGAVVMVLMVFN 480
QY 481 AVMAKTKTYOVANHKSKDNRIKLMNEILNGIKVLKLYAMELAFKXVLAIROBELKVLK 540
D 481 AVMAKTKTYOVANHKSKDNRIKLMNEILNGIKVLKLYAMELAFKXVLAIROBELKVLK 540
QY 541 KSAIYLSAVGTFTWCTPFLVALCTPAVYVTTDENNIIIDAQAFVSLALFNIRFPLNITP 600
D 541 KSAIYLSAVGTFTWCTPFLVALCTPAVYVTTDENNIIIDAQAFVSLALFNIRFPLNITP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEBLEPDSIERRPVKDGGTNSIVRNATFTWASDPT 660
D 601 MVISSIVQASVSLKRLRIFLSHEBLEPDSIERRPVKDGGTNSIVRNATFTWASDPT 660
QY 661 LMGITFSIPEGALVAVVQVQGGKSSLLSALLAEMDKYEGHVAIKGSVAVYVPOQAMIOND 720
D 661 LMGITFSIPEGALVAVVQVQGGKSSLLSALLAEMDKYEGHVAIKGSVAVYVPOQAMIOND 720
QY 721 SLRENILFEGCQLEBYRHSVIOACALLPDLLETBSGDRTEIGEKVNSGGOKORVSLAR 780
D 721 SLRENILFEGCQLEBYRHSVIOACALLPDLLETBSGDRTEIGEKVNSGGOKORVSLAR 780
QY 781 AVYSNADYYLFPDDPLSAVDAAHVKHIFENVIGPKGMLKNKRIILVTHSMSYLPQVDVITV 840
D 781 AVYSNADYYLFPDDPLSAVDAAHVKHIFENVIGPKGMLKNKRIILVTHSMSYLPQVDVITV 840
QY 841 MSGGKISEMGSYOELLARDGAFAEPLRTYASTEQDABENGVTGVSQGEKAKOMENG 900
D 841 MSGGKISEMGSYOELLARDGAFAEPLRTYASTEQDABENGVTGVSQGEKAKOMENG 900
QY 901 LVTSBAGQOLROULSSSSSYSGDISRHNSITAELOKAEKKETWKLMEADKAQTGOYKL 960
D 901 LVTSBAGQOLROULSSSSSYSGDISRHNSITAELOKAEKKETWKLMEADKAQTGOYKL 960
QY 961 SVYMDYMAKIGLFTSFLSIPLFCMCHVASLANSNYMLSLMTDPIYNGOEHKVALSYG 1020
D 961 SVYMDYMAKIGLFTSFLSIPLFCMCHVASLANSNYMLSLMTDPIYNGOEHKVALSYG 1020
QY 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLHSLILNSPMSPFERTSGNLYNFSKEL 1080
D 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLHSLILNSPMSPFERTSGNLYNFSKEL 1080
QY 1081 DTVDMSIMEVITKMGSLFNVIAGCTVILLATPIAIIIPGLIYFPVQRFYVASSROL 1140
D 1081 DTVDMSIMEVITKMGSLFNVIAGCTVILLATPIAIIIPGLIYFPVQRFYVASSROL 1140
QY 1141 KRLESVSSPYVSHNETLLGVSVIRAFEEORFIHQSDLKVDENOKAYPSIVANRWLA 1200
D 1141 KRLESVSSPYVSHNETLLGVSVIRAFEEORFIHQSDLKVDENOKAYPSIVANRWLA 1200

```

```

QY 1201 VRLCEVGNCTVLPALPAVIRSRHSLAGLVGSYSISLOTTYYINMLVRSSSEMETNIVA 1260
D 1201 VRLCEVGNCTVLPALPAVIRSRHSLAGLVGSYSISLOTTYYINMLVRSSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVAFRYCRLRYRDLDFVLVSHINVTINGG 1320
D 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVAFRYCRLRYRDLDFVLVSHINVTINGG 1320
QY 1321 EKVIVGRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKITIIPQDPVLF 1380
D 1321 EKVIVGRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKITIIPQDPVLF 1380
QY 1381 SGLRPMNLDPSSQYSDBEWVTSLELAHKDFVSAIPDKLDHECAGGENSVSGOROLVCL 1440
D 1381 SGLRPMNLDPSSQYSDBEWVTSLELAHKDFVSAIPDKLDHECAGGENSVSGOROLVCL 1440
QY 1441 ARALLRKTLIVDEATPAVDLEFTDDIIOSTIRFOFEDCTVLTARHLNTIMDYTRVVL 1500
D 1441 ARALLRKTLIVDEATPAVDLEFTDDIIOSTIRFOFEDCTVLTARHLNTIMDYTRVVL 1500
QY 1501 DKGEIOEYGAPSDILLQORGLFYSMAKDAGLV 1531
D 1501 DKGEIOEYGAPSDILLQORGLFYSMAKDAGLV 1531

RESULT 2
JE0336
canalicular multispecific organic anion transporter - human
C:Species: Homo sapiens (man)
C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004
C:Accession: JE0336
R:Uchiyama, T.; Hinochita, E.; Haga, S.; Nakamura, T.; Tanaka, T.; Tsh, S.; Furukawa, M.;
Biochem. Biophys. Res. Commun. 252, 103-110, 1998
A:Title: Isolation of a novel human canalicular multispecific organic anion transporter,
t.
A:Reference number: JE0336; MUID:99032812; PMID:9813153
A:Accession: JE0336
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1527 <UCH>
A:Cross-references: UNIPROT:O15438; UNIPARC:UPI0000169859; GB:AF083552
C:Superfamily: human multidrug resistance protein cMOA12; ATP-binding cassette homology
C:Keywords: ATP
F:1306-1499/Domain: ATP-binding cassette homology <ABC2>

Query Match 45.3%; Score 4484.5; DB 2; Length 1527;
Best Local Similarity 56.6%; Pred. No. 4.7e-296;
Matches 871; Conservative 261; Mismatches 372; Indels 35; Gaps 9;

QY 8 SADSDPLMDMNVMTNTPNDFTKCFONTVLVWVPCFYLMACFPFYLYLSRHDGYIQM 67
D 8 SADSDPLMDMNVMTNTPNDFTKCFONTVLVWVPCFYLMACFPFYLYLSRHDGYIQM 67
QY 67 SHLSLKNVLDVLLMVCWSMADLFYSFHGLVGRAPAPVFTPIPVVGVTMLATLTIQYE 126
D 67 SHLSLKNVLDVLLMVCWSMADLFYSFHGLVGRAPAPVFTPIPVVGVTMLATLTIQYE 126
QY 128 RRGVQSSGIMLTFLMVALYCALAIIIRSKIMTALKEDAQVDLFEDITFYVYSLLIQV 187
D 128 RRGVQSSGIMLTFLMVALYCALAIIIRSKIMTALKEDAQVDLFEDITFYVYSLLIQV 187
QY 187 LSCSDRSPLFSETIHDNPPCBSSASFLSIITFWITGLIVRGYRQPLESDIWSLNKE 247
D 187 LSCSDRSPLFSETIHDNPPCBSSASFLSIITFWITGLIVRGYRQPLESDIWSLNKE 247
QY 247 DRSGVTVQQLLEARKKQKOTARH-----KASAPGK-----NASGEDEVLTGARDP 294
D 247 DRSGVTVQQLLEARKKQKOTARH-----KASAPGK-----NASGEDEVLTGARDP 294
QY 308 EMNBSLFPVLVKTGFPYFLMSFFPKAIIHDLMMFSGPOLIKLIRKVNNTKAPDMQGYRYT 367
D 308 EMNBSLFPVLVKTGFPYFLMSFFPKAIIHDLMMFSGPOLIKLIRKVNNTKAPDMQGYRYT 367
QY 367 R-KSFTLLATRGSSFLISACFKLIDLLSFLNPDLSILIRISNMPGSMWGFVLA 353
D 367 R-KSFTLLATRGSSFLISACFKLIDLLSFLNPDLSILIRISNMPGSMWGFVLA 353

```

368 VLEFVTAQLQTLVHOFHICFVSGMRITKAVIGAVREKALVITNSARKSTVEIYVLM 427
354 GIMELCSMMGSLILQHYHYHIFVTVGKRTGIMGYIRKALVITNSVRSATVEIYVLM 413
428 SVDAQRFMDLATYINMTWSAPLQVILALYILMLNGPSVLAVGAVMVLMEVNAVMMKMT 487
414 SVDAQRFMDLATYINMTWSAPLQVILALYILMLNGPSVLAVGAVMVLMEVNAVMMKMT 473
488 KTVAVAHMKSMDNLIKMLNELNGIKYIKUYAMELAFDKYLAIRQELKYLKKSAYLSA 547
474 RAFCVQKQKDKSDIKMSIELNGIKYIKUYAMELAFDKYLAIRQELKYLKKSAYLSA 533
548 VGTFTWVCTPFLVACTFAVAVVITDENNIIDAQTAFAVSLALENLIRPILNIPWISIV 607
534 TTTFTWVCTPFLVACTFAVAVVITDENNIIDAQTAFAVSLALENLIRPILNIPWISIV 593
608 QASVSLKRLIFLSHEELPDSIRRRPYKDGGSITVRNATFTWASDPPTNGITFS 667
594 QASVSLKRLIFLSHEELPDSIRRRPYKDGGSITVRNATFTWASDPPTNGITFS 650
668 IPEBALVAVVQVCCGKSSLSLALAEKDKVEGHALKGSVAVYPOQAMIONDSLRNLT 727
651 VPKGALVAVVQVCCGKSSLSLALAEKDKVEGHALKGSVAVYPOQAMIONDSLRNLT 710
728 FGQCLSEPPYRSVIOACALPDLIELPSGDRTEIGEKGVNLGGOKORVSLARAVSNAD 787
711 FGKALNRKRYOQTEACALLADLEMLPGDOTEIGEKGVNLGGOKORVSLARAVSNAD 770
788 IYLFDDPLSAVDAAVHKHIFENVIQPKGMLKNKTRILVTHSMXYLPQVAVIYVSGKIS 847
771 IFLDDELPSAVDSHVAKHIFDHVIGPEGVLAKTRIVLVTHGISLPLQDPFIVLADQVS 830
848 EMGYSQELIARDAFAEFLRTYASTEOBQAEENGCVTVSGPGEAKOMENGM----- 901
831 EMGYPPLALORNGSFANFLCNVADDEQGHLEMTALKEADGKALIEBTLNHNHDLT 890
902 ----VTDSAGKOLOROSSSSSYSGD-----ISRHNHSTAE-LQKAEKKEETMKLEAD 951
891 DNDPVTYVQKQFQRKQSAUSS-DGEGQGRVPRRHIGPSKQVYTEKADGA--LTOEE 947
952 KAQGVKLSVYMDYMKALIGLIFSLIFLFCMCHVSLASNMVLSMTDPIYNGTOEH 1011
948 KALIGVELSVFMDYAKAVGLCTTLALCLLVGQSAALIGAVMLAMTNDAMDSDQN 1007
1012 TKVRLSYGALIGISQGIAYFGYSMAVSIQGITLARSCLHDLHLSILSPMSFERTPSGN 1071
1008 TSIRLGYAALIGIQQFLVYMLAAMAAAGCIGQARVHLQALHNKIRSPSFFPTPSGR 1067
1072 LVNRFSEKELDTVDSEMIPEVIMFMGSLFNVIGACIVILLAPLAIILIPGLIYFFVOR 1131
1068 ILNCFSKDIYVDEVLAPVIMLNLNSFFNALSTLVIMASTPLTVVILPLAVLYTLVOR 1127
1132 FYVASSQOLKRLSESVSRSPVSHFNETLLGVSIVIRAEBOERFIHQSDLKDEQKAYYP 1191
1128 FYAATSRQOLKRLSESVSRSPVSHFNETLLGVSIVIRAEBOERFIHQSDLKDEQKAYYP 1187
1192 SIYANRLAVLBEVCNACIVLPAALPAVISRHSLSAGLVGLSVSYLQVTTYINMLVMS 1251
1188 YIISNRMISIGVEFVNCVVLPAALPAVISRHSLSAGLVGLSVSYLQVTTYINMLVMS 1247
1252 SEMETNIVAVERLKEYSETEAEAPMOQOETAPPSMPOVQGRVERRANCCLARYRELDVLR 1311
1248 SDLESNTVAVERVEXKTEAEAPVVEGSRPPGMPRGEVEFRNRSVRRPDLIVLR 1307
1312 HINVTINGEKVIGRTGAGKSSITLGLFRLNESAAGEIILIDGINAKIGLHDLRFKIT 1371
1308 DLSLHVHGEKVGIVRTGAGKSSITLGLFRLNESAAGEIILIDGINAKIGLHDLRFKIT 1367
1372 IIPDDPLFSGSLAMNDPFSQYSDSEVWISLELAHLKDFVSLAPDLKHECAGGENLS 1431
1368 IIPDDPLFSGSLAMNDPFSQYSDSEVWISLELAHLKDFVSLAPDLKHECAGGENLS 1427
1432 VGGQVCLARALLRKTRILVLEDEATVADLETDDLQSTIRTRFEDQVYLTIAHRLNTI 1491

1428 VGGQVCLARALLRKTRILVLEDEATVADLETDDLQSTIRTRFEDQVYLTIAHRLNTI 1487
1492 MDYTRVVLVDKGEIQEYGAQSDLLQQRGLYSNAKADGL 1530
1488 MDYTRVVLVDKGEIQEYGAQSDLLQQRGLYSNAKADGL 1526

RESULT 3
S71841
multidrug resistance protein, canalicular - human
C/Species: Homo sapiens (man)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
R/Accession: S71841, S71840
R/Koenig, J.; Keppler, D.
submitted to the EMBL Data Library, August 1996
A/Reference number: S71841
A/Accession: S71841
A/Molecule type: mRNA
A/Residues: 1-1545 <KOE>
A/Cross-references: UNIPARC:UPI00001746CD; EMBL:X96395; NID:g1507819; PIDN:CAA65259.1;
R/Buechler, M.; Koenig, J.; Brom, M.; Kattebeck, J.; Spring, H.; Horle, T.; Keppler, D.
J. Biol. Chem. 271, 15091-15098, 1996
A/Title: cDNA cloning of the hepatocyte canalicular isoform of the multidrug resistance
A/Reference number: S71839; MUID:96279006; PMID:8662992
A/Accession: S71840
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1415-1429, 'VP', 1432-1455, 'E', 1457-1545 <BUE>
A/Cross-references: UNIPARC:UPI00001746CE; EMBL:X96395
A/Genetics:
A/Genes: GDB:ABCC2; CMOAT; ABC; MRP2; CMRP; DJS
A/Cross-references: GDB:6089489; OMIM:601107
A/Map position: 10q24-10q24
A/Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
A/Keywords: ATP; glycoprotein; nucleotide binding; P-loop; transmembrane protein
F/654-837/Domain: ATP-binding cassette motif A (P-loop)
F/671-678/Region: nucleotide-binding motif A (P-loop)
F/1317-1510/Domain: ATP-binding cassette homology <ABC2>
F/1334-1341/Region: nucleotide-binding motif A (P-loop)

Query Match 36.2% Score 3587; DB 1; Length 1545;
Best Local Similarity 47.5% Pred. No. 5.3e-215;
Matches 746; Conservative 286; Mismatches 461; Indels 78; Gaps 20;

3 LRFGCSADGSDPLMDNVNTNTS-----NDPFTKCFONTLVVWPCEYLWACPEPFL-- 55
2 LEKFCN-----STFNSSFLDSPEADLPFCFQYTLVWVILGLFWLWLPWQGLHV 51
56 YLSRHDRGYIQMTPLNKTKTALGFLMYCVALDFYSFWERSKGIPLAPVPLVSPITLGI 115
52 YKSRITKRSSTTKLYLAK-QVFGVFL-ILAIETALVLTEDSQCATPAVARYTNPSLY-L 108
116 TLLATFLIOLERKGVSSGIMLT--FWIALVCAALILRSKIMTALKEAQVDLPDIT 174
109 GTWLVLLIYSSQWOCQKSWFLSFWILSLICGTFQFOTLIRTLQSG-NSNLAYSL 167
175 FYVYFSLILQVLSGFSRPLPFSRTIHDPNCPBSSASFLSRITFMWITGILVNGYRQ 234
168 FFIISYQRIILILFSAFSEN-----NSSNNPSSIASFISITYSWSDSIILIKYKR 219
235 PLEGSIMLSINKEDTSEQVPLVVKWKECAKTRK-----QPVKVYSSKD 281
220 PLTLEVYMEVDEEMKRTKTLVSKETEMKRELQKARALQRRQKSSQONSQALPGLINK 279
282 PAQPKSSKVDANEEYBALIVKSPKEMNNS--LFVLYVTGEPYFLMSPFFRAIHDLM 339
280 QSGSODALV---EDVEKKKSGYKQDVPSKMLKALFTFYMWILKSLFLKLVNDITF 336
340 FSGPQILKILIKFVNDTKADWQGYFYVTLQTLVLAHQYHICFVSGMRITKAV 399
337 FVSPQILKILISFASBRDTYLMIGVLCALILFLFALIIQSFCLQCYQLCFKLDGVKRTAI 396


```
Db 344 GFVSSNSYVWFYICALIMFAVTLIQSFCLQSYFOHCFVLGCMCVRITVMSIYKALTL 403
Qy 411 TNSARKSTVEIYNLMSVDQRFMDLATYINMISAPLOYITLALYLMNLGSSVLAV 470
Db 404 SNLRKQYITGETYNLMSVDSQKLMDATYIMQLVMSVYIQTTLIFPLMRLGPSILAGV 463
Qy 471 AVMYLVNVAVMAMKTKTYQVAHMSKDNRIKLMNELINGIKYKLYAMELAFKDYLA 530
Db 464 GVMVLLIPYNGVLTAKIRNIOVMKDKRLKIMNELISIKIKLKPFAMSPSQEOYQG 523
Qy 531 IROELKVLKKSAYLSAVGTFTWVCTPFLVALCTFAVYVTIDENNIIDAQTAFAVSLALFN 590
Db 524 IRKKEKNTLRFQQLSLIFILQITPLVSVTFVSYVLVDANSVNLAKAFSITLLEN 583
Qy 591 ILRPLMLPWTSSIIQASVSVDRLEBYLGGDDLDTSIRRVSNFD---KAVKFSBAS 639
Db 584 ILRPLMLPWTSSIIQASVSVDRLEBYLGGDDLDTSIRRVSNFD---KAVKFSBAS 639
Qy 651 FTNARSDPPTLNGITFSIPEGALVAVYGVGCGKSSLSALAEKMDYEGHVAIKGSVAY 710
Db 640 FTNARSDPPTLNGITFSIPEGALVAVYGVGCGKSSLSALAEKMDYEGHVAIKGSVAY 710
Qy 711 VPQQAWTQNDLRENILFGCOLBEPYKSVIQAQALLPDLIELPSGRTIEGKGVNLG 770
Db 700 VPQQAWTQNDLRENILFGCOLBEPYKSVIQAQALLPDLIELPSGRTIEGKGVNLG 770
Qy 771 GQKQSVLARAIVSNADITYLFDPLSAVDHVGKHEBNVIGPKYMLKNTKRLIVTWSMS 830
Db 760 GQKQSVLARAIVSNADITYLFDPLSAVDHVGKHEBNVIGPKYMLKNTKRLIVTWSMS 830
Qy 831 YLPQVDIYVMSGGKISEMGSYQELLARDGAFAEFLPTYS-TQEOADAENQVTVGSGP 889
Db 820 FLRPFVDEIVVLGKTLLEKSGYRDLKKGVFARNMKTIPMKHSGPEEGATVNN----- 872
Qy 890 GKEAKQWENGLVT-----DSAG-----KQIQROLSSSSSSYGDISRHNNSTAEIQ-- 935
Db 873 DSEAEDEDDGGLPMEBIPEDASLARMRENSLARTISRSSRSSRKSLAKNLTAKNV 932
Qy 936 ---KAQKKEBTWKLMDAKAQIQGVKLSVYWDYMKALGFLISFLFENCNHVALAS 992
Db 933 NVLKEKEKEVEGQKLIKKEFEVETGKVFSTYKLTQAVGWSLIFILLFYGLNNVAIGS 992
Qy 993 NYWMLSTLT-DDEPIYNGT---QEHKRVLSVYGAIGISOGIAPGVSAVSIIGILLARCL 1048
Db 993 NYWMLSTLT-DDEPIYNGT---QEHKRVLSVYGAIGISOGIAPGVSAVSIIGILLARCL 1048
Qy 1049 HVDLHSLRSPMSFPERTPSGNLVNRFSKELDTVDSEMIPEVIMKFGSLEFNIGACIVI 1108
Db 1053 HGQLLTNITLAPRPFPTTPTPIRIVNRFSGDISIVDDLLPQTLASMMWCFPGIAGTIVMI 1112
Qy 1109 LLATPIAIIIPYGLIYFVQRYVASSRQLKLESVSNPVSHEFNELLGVSIVIRAF 1168
Db 1113 CMATPVAIIIPYGLIYFVQRYVASSRQLKLESVSNPVSHEFNELLGVSIVIRAF 1168
Qy 1169 EBOGRFTHODLAKDENOKAYRBSIVANRMLAVLLECNGCIVLFAALLPAVISHHSIAG 1228
Db 1173 EBOGRFTHODLAKDENOKAYRBSIVANRMLAVLLECNGCIVLFAALLPAVISHHSIAG 1228
Qy 1229 LVGLSVSYSLQVTTYLMLVMSSEMETNIVAVERLKEYSETEKEAPMOJQETAPSPSP 1288
Db 1233 LVGLSVSYSLQVTTYLMLVMSSEMETNIVAVERLKEYSETEKEAPMOJQETAPSPSP 1288
Qy 1289 QVGRVERRATCLARTRELDPLVLRHINVTINGEKVYIVGRTGAKSSLLGLGRINESAE 1348
Db 1292 RHGSIQNNYQVRRPELDLVLKGTICNINSGEKVYIVGRTGAKSSLLGLGRINESAE 1351
Qy 1349 GEIITIDINTAKIGHLRFKTIITIPDDPVLFSGSLAMNLDPSQVDEDEWMSLELAHL 1408
Db 1352 GQIITIDINTAKIGHLRFKTIITIPDDPVLFSGSLAMNLDPSQVDEDEWMSLELAHL 1411
Qy 1409 KDFVSAIPDKLDHCBAGGENTSVGORQLVCLARALLRKTKIILVLDATAVAVDLETDLI 1468
Db 1409 KDFVSAIPDKLDHCBAGGENTSVGORQLVCLARALLRKTKIILVLDATAVAVDLETDLI 1468
```

```
Db 1412 RSFVSGQLGLLSETEGDNLSIGRQLTCLGRAVLRRKSKILVLDATAVAVDLETDLI 1471
Qy 1469 OSTIRQPEDQCYVTLTIRNTIMDTRVVLVDXGEIOEGAPSDLLQOGLFYSMAKDA 1528
Db 1472 QTYIRKESQCYVTLTIRNTIMDTRVVLVDXGEIOEGAPSDLLQOGLFYSMAKDA 1531
Qy 1529 GL 1530
Db 1532 GI 1533

RESULT 5
E89447
protein F57C12.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence Revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: E89447
R:anonymous, The C. elegans Sequencing Consortium.
Science 283, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C.el
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an
A:Accession: E89447
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1494 <STO>
A:Cross-references: UNIPROT:Q20943; UNIPARC:UP1000017801E; GB:chr_X; PIDN:AAA83299.1; P
C:Genetics:
A:Gene: F57C12.4
A:Map position: X
C:Superfamily: human multidrug resistance protein CMOAT2; ATP-binding cassette homology

Query Match 33.2%; Score 3289; DB 2; Length 1494;
Best local Similarity 47.6%; Pred. No. 9.3e-215;
Matches 696; Conservative 239; Mismatches 435; Indels 92; Gaps 18;

Db 119 LAFPLQLERKQVSSGIMLTFWLVALVCAALILRSKINTALKEDQVDFDITRYVY 178
Db 76 LAILTVACNKKKITTSGVTLTWLVVVGIGIEFFRYLSGFIYNEVALGIRATLVIA 135
Qy 179 FSLILQLVLSCEFSRSPLESETIHDNPPCESSASFLSITFWITGLIVRGYRPLEG 238
Db 136 FTSALALEFLCCRAD---VSDMYKSSSCPEYTAGSINLTQWFTGLAVLNKSKLEN 192
Qy 239 SDIWSLUNKEDTSQVQVAVLVKMKKCECAKTRKQPVKYVSSKDP--AQPKSSKVDANEE 296
Db 193 EDLMDINEIDKAENLIPSFQNLKPRIDEYH-QNIK-----KDBSALPKN----- 237
Qy 297 VEAIVKSPKQENNPISFKLYTYTTPGYFLMSFFPRKIHDLMMFSGQIILKLIKPVNDT 356
Db 238 -----HPSFVPIPFYTYKTYTLLAGFFYKLCFMDQFLAPQLLKOLIGFIEDK 284
Qy 357 KAPDMQGYFYVLLPFTACQLQTLVLRHQYFHCIVSGMRKITAIVAGYAKALVITNSARK 416
Db 285 NQVWVIGCSIVGIMFSSFLQSMFLHQYISMRLGNHVASVLTSAVYSALMINSERAK 344
Qy 417 SSTVGEIYNLMSVDAQRFMDLATYINMISAPLOYITLALYLMNLGSSVLAVAVVLM 476
Db 345 GKTIQAIYNLMSVDAQRFMDLATYINMISAPLOYITLALYLMNLGSSVLAVAVVLM 476
Qy 477 VPRNVAVMAMKTKTYQVAHMSKDNRIKLMNELINGIKYKLYAMELAFKDYLAIROEL 536
Db 405 LPVNGIATQMRCCQEQEMKDKBERIKMSEILNMGKVLKLYSMERSMEMNVKIRIREL 464
Qy 537 KVLKSAVYSVAVTFTWVCTPFLVALCTFAVYVTID-ENNIIDAQTAFAVSLALFNIRP 595
Db 465 HIILKLSYFPAALVFSMICKAPLASVSVYVYVLDPENNVLTPELTVALSLFDILRMP 524
Qy 596 LNLIPVYISSIVQASVSLKRLIFLSHELEPDSIRRPYKDGGSITVNNATFTWAR 655
Db 525 LAMVAVYGBAVQCSVSNTRLKEFFAAEWSPQT---SISHGETSALIEVENGLFSWSS 580
```

Qy	566	SDPPLNNTITSIPGALVAVVGVGCGSSLLTSLTLMEMDVEBHVAKSVAIVPOOA	715
Dp	581	DEDPILBRSIKRIQKGQIVAIYVKGKSGKSSLLHLLBEMKSSSVQINGIAIVPOOA	640
Qy	716	WIONDSLENTILFGCQLEBEPYRYSVIAQCALLPDLEILPSGDRTIEIGKGVNLSGQOKOR	775
Dp	641	WIONMSLENNILFNKPYULENVEDVVKNCALKEEDLANIPAGDRTIEIGKGINLSGQOKOR	700
Qy	776	VSLRAVYSNADIIYLPDDPLSAVDNAHVGKHFENYIGRK-GMLKXKTRILVTHMSYLPQ	834
Dp	701	VSLRAVYQNPDIILDDPLSAVDNHVGKHFENYISSGTGCLAKTRVLVTHGTYLKH	760
Qy	835	VDVIIWMSGKISSEMGSYOELLARAGAFIEPLRTASTPEQOEDBAENGTVGSGPG----	890
Dp	761	CDQILVLEKGTISSELTGYOELLNNSGAFIEPLLEBFLIEBSKTRGVASTIGDGSVDEIL	820
Qy	891	KEAKOMENGLVTSAGQOLOROLSSSS-----SYSGDISRH-----HNSTAELOK	936
Dp	821	RDLOGVKRGIL-----KRLBSHLSQESDKEPDSARATIEYSDSSRRSVLHSPSGHBE	874
Qy	937	AEA-----KKEETWKLMEADQAQGGQVLSYYWDMYKAIGLFISFLSIFLFCMHV	987
Dp	875	NEALIGASIEDVPACENTOLIEKEFVEKGVKFEYIYVFOAISIPITLLPFLFVYGSAG	934
Qy	988	SALASNNYMLSMITDO-PIVNGTQHTKYRLSYVAGLAGISQIAGVGSAAVSGIGILASR	1046
Dp	935	LGIISNFIPLAKISDAKSGNRTSSPAKELGIYAVLGMQSVVLIASTIILITGVLRASR	994
Qy	1047	CLHVDLHSLIRSPPMSFEFERTPSGNLVNRFSELDTVDSMIEVIMKFMGSLFNVTGACT	1106
Dp	995	ILHAGLILNIMKSPAPAFDVPITPGRILNRRIGDIEBRLTDPVIRHMSMTIFNVAVTLV	105
Qy	1107	VILLATPLAIIIPPLGLIFPVQRRFYAASSQOLKRLBSVSSPVYSHNPELLEGVSYR	1166
Dp	1055	VIMATTPWAGIAFALISVYFVLEFXYISTSRQOLKRLBSASRSPYISHPOESIOGASAIR	1114
Qy	1167	APEEGERFPHOSDLDVENDOKRYYSISYANRMLAVRLECVGCIYU.PAL.FAVISRHS--	1222
Dp	1115	AFGVVDNFIKOSQORVDDHLAIYFYSIYANRMLAVRLEKVGMLVLSAAGAAVYRDEBFG	1177
Qy	1225	LSAGLVGVSYSLOVTTYLANLVMSSEMENTNIAVERLKEYSBTEKEAPMOIETAPP	1284
Dp	1175	LSAGLVGVSYSYALNITQTLNNAVAMTSELENNIYSEVLEIKXYTYPITPEG--NNSRRLAA	123
Qy	1285	SSWPOVGEVEFRNYCLRTREDDLVLRHINVTINGEKYGIYVGTGAKSGLTLGFRIN	1344
Dp	1233	KSWPKGEISINPFVRYRPGDILVHGISAHIAISEKGIYIGRTGAKSGKSLTLALFRIT	1292
Qy	1345	ESABEELIINDGININAKIGHDIRFKITIIPODPV.FSGSLRNNLDPFQSYGSEFWTSL	1400
Dp	1293	EADGSGSIEIDGININLQDEQARSCTITVPDPVLPFSGTKRNNLDPFSYISOSYWEAL	135
Qy	1405	LAHLDPFYSALPDKLDHCECAGEGNL-----SYGORQLVCLALALRLKTYI	1455
Dp	1353	NAHLKPFYKSLQDGLBHKISBEGENULRPAKSGSKQRFQLNFSVGGQQLICLAALLRKTYV	1411
Qy	1451	LVLDATAVADLETDDLIOSTIRTOPEDCTVLTIAHRLNTINDYTRVIVLDKGEIOEYGA	1510
Dp	1413	LVLDATAAIVADVETSLIOKTRIEQFKECTVLTIAHRLNTVMSDRLLVLIDKGRAVEFDS	1472
Qy	1511	PSDIL-QORGLFYMAKDAGLV 1531	
Dp	1473	PNLLANPDGIFYMAKDAANV 1494	

RESULT 6
T21219
hypochemical protein F21G4.2 - *Caenorhabditis elegans*
CISpecies: *Caenorhabditis elegans*
CDate: 15-Oct-1999 #sequence_revision 15-Oct-1999 #next_change 09-Jul-2004
CAccession: T21219, T24002
RIMortimore, B.
submitted to the EMBL Data Library, October 1996

A:Reference number: Z19392
 A:Accession: T21219
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1573 <WT>
 A:Cross-references: UNIPROT:Q93552; UNIPARC:UPI00000823F7; EMBL:Z81016; PIDN:CAB02667.1
 A:Experimental source: clone F21G4
 R:Kershaw, U.
 Submitted to the EMBL Data Library, March 1995
 A:Reference number: Z19828
 A:Accession: T24002
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1573 <WT>
 A:Cross-references: UNIPARC:UPI00000823F7; EMBL:Z48621; PIDN:CAA8549.1; GSPDB:GN00028;
 A:Experimental source: clone R07B1
 C:Genetics:
 A:Gene: CESP:F21G4.2
 A:Map position: X
 A:Introns: 21/3; 122/3; 197/3; 276/1; 654/1; 911/1; 1067/1; 1472/2; 1537/2
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
 Query Match 32.6%; Score 3225; DB 2; Length 1573;
 Best Local Similarity 42.4%; Pred. No. 2,3e-210;
 Matches 685; Conservative 305; Mismatches 490; Indels 134; Gaps 26;

```

0Y 3 LRG---FCSAGSDJLMMNMVNTVNTSN--BDKFCQONTLVVWPCEYLWACFPFYLYLS 58
Db 9 LKGIADFC-----GERYMD-PAYMNASYPIILSQOCHTLVWFP----- 47
0Y 59 RHDRGYIOMTPLNKTXTALGFLMLIVCMADLFYSF-----WER-----SR 98
Db 48 -----TAIVFLLAPILTAQOIFRRPNPIMPRRRIOLKIGLACILYADAL 91
0Y 99 GFLFAPVF-----LVSPFLLGITTLTALFQLERKQVSSGIMLFWLVAL 146
Db 92 SLFTVAIETTLFQGPFAVDVFPYFLLCLAMVVALIIVSCNNYGLVTSGGFLFISHLVPT 151
0Y 147 VCAIAIIR---SKIMTALKEDAQVDLFRDITFVYVYFSSLLIQLVISCFSRSPFL-FSEIT 2022
Db 152 ISAIPELLIYMIQOIVNPAEAMWIDYPRCIAFIWFECAPETYLHCYADASPEGYKLS 211
0Y 203 HDPNRCPPSSASFLSRTIFPMWITGLIVNGYRPLBEGSDIWSLUNKEDTSQVYVULVKNK 262
Db 212 AARNSPETTSSEFLNRTIMWMMNPNSSCSGKVPKPLEVSDLSYSEADTSLNLVPMKYNLMD 271
0Y 263 KECATRRKOPVAVVYSSKDPAPKPESSKVDANEVEAL-----IYKSPKQEWN 310
Db 272 KQSKKFEETAARRRGSN--ASRTNRRRRRSSNDTTPILNDQSTDYDGVSPAGQSTOK--M 327
0Y 311 PSLFVLYKTCGPYPLMSEFFPAIHDLMMFSGPOLIKLICKVNDTKADPMQGYFTVLL 370
Db 328 PSIIITLFLMEFMDVITAMFVLLDSVLLFCNPLBLKLSIRFTEBERPMQGVLAFTM 387
0Y 371 FYTACLOTLLVHQYHICFVSGMARIKTAVIGAVYRKALVITSARKSIVGEIVNLMSVD 430
Db 388 FFSALSTSLISHYILMTKRVGSTRQOTCLTAAYIRKTLRLSNAREKTVGELVINLMALD 447
0Y 431 AORFMDLATYIMMISAPLOVILALYLLMLNLGSPVLAVGAVVMVMPVNAVAMKTKTY 490
Db 448 VDRFOQITPQTMQYMSNPFQIGLAFLLPFOQGVSVFSVAVVMLLFPINFIVTMIRKW 507
0Y 491 QVAHMKSDNRKJLMMELINGIKVLKIVAMELAFKDXLAIROEELKVLKKSAYVLSAVGT 550
Db 508 QIAQMYTYDERKXKMNVEVINGIKVLKIVAMEPMEGVIEDLROEELIKKAFLETFSD 567
0Y 551 FTWVCTPFLVALCTFAVVYITDENNIIIDQAQTFVSIATFNIIIRFLNIIIPMVSSIVQAS 610
Db 568 MUMTASPFVIALSTFATFIYIDPKVILPEIAFVSLTFLNOLRSRPSQVABELITQVQVY 627
0Y 611 VSLKRLRIFLSHEELPEPDSIERRPVKYDGGGNTSITVRATFTW--ARSDP-PLINGITFS 667
Db 628 VSNRIKKEVLSEELNVEAIDHR-ARD--NDNVICTIKEKCLSMESAEHOPVPLTINISFS 684

```

```
Qy IPEGALVAVGVGGGKSSLLSALLAEMDKVEGHVIAIKGVAVYVPOQAMIONDSIRENLT 727
Dy VNRGQLVTVVGVGGKSSMLQALMGEMEKLSGSIIMGRICVYPOQPMONNITRQNT 744
Qy 728 FGCCLDEEYRVSIVQACALLPDLLEITPSGDRTEIGEKVNLGGCKQKVSILARAVYNSAD 787
Dy 745 FGKQFDEYFGRVADACALYRDLOILPLGDNTEIGEKINISGGCKAISIARAVYQNDH 804
Qy 788 IYLEDDEPLSANDAHVKHIFENVIGPKMKLNKTRILVTHSMSTYLPQVDVIIWVGSKIS 847
Dy 805 IYLLDDPMSAVDAHVSQIFGSGVIGPEQMKNKTRILVTHNEISFEKSDLLIVNMGKIE 864
Qy 848 EMGSYQELLARDGAFAEFL-----RTYASTQEQDABENGVTVSGGCKAKQKQEN 898
Dy 865 YSGKTDIDM-QQGAPEQLLIECEKEERREKREASADSDENSBEGIMIGDSDFEYDD 923
Qy 899 GMLVDSAGKQLOKOLSSS--SSYSGDISRHNNST--ELQKAKEKEETWKLMEA--- 950
Dy 924 DVM-----ASPIIDHVLGTHSHSTVSGIINRRRISTSTHKORRLSTTSHSTHSTASTQ 979
Qy 951 -----DKAQTOYKLSVWDYMKAGLFTSPLSTFEMCNHVSALASNTWLSLWTD 1003
Dy 980 TROLGTERVETGKMDTYKKYFGAMGMSIALVFLQMTTSTJFSMGRNMLTDMSDN 1039
Qy 1004 IV-----NOTQHTKRLSVYALGISOIAVFGSMANVIGGILASRCILAVDLHSITRS 1059
Dy 1040 AARGSNITGQPIAIRLGVAVGLPSEIILFICMLSTLYGGVSNRLAHPLKRNLF 1099
Qy 1060 PMSEFETPSGNLVNRFESKELDVDSMIPEVIKMFMSGLFNVIGACIYLLATPLAII 1119
Dy 1100 PMAYYDTTPFRILNRICKDLEIYDVLPRVQFPACQLLOVSTLIIMSTYFVGI 1159
Qy 1120 PPLGLIYFVQRFYVASSRQKLESVRSRPSVSHFNMTLLGVSVIRAFEBQERFIHQSD 1179
Dy 1160 IPLSVMLVMVRYYIATSROKLESITRSPYSHLSISIGSATIRAYHLDVDFCKLSE 1219
Qy 1180 LKUDENQKAYPSIVANRMVLRLECVGNCLVPLALPVISSRLSAGLVGSYSLSQ 1239
Dy 1220 TKVDSHVQCRFLYANVARMVSRLEFGNCIVLPSALFALTRITTTISGVGLSVYALN 1279
Qy 1240 VTYTLNMLVRRMSMETNIVAVERLKEYSETEKEAPMOIE-TAPSSMPQGVVEPRNY 1298
Dy 1280 ITTYLVANVROITKLEINIVSEVKEYAETETAEMKSEPKRPPQWPEBEGIVMANY 1339
Qy 1299 CLRYREDLDPVLRHINVTINGEKEVGIAGRTGAKSSLTLLGFRINSABEIIIDGINI 1358
Dy 1340 SARVRPGINLVVQOLNVEIKPHEKGIYGRGAGKSVTSLFRIIEAABEQIVVDGINL 1399
Qy 1359 AKIGLHLDRFKITITIPQDPVLFSSSLRNLDLPPSOYSPEEYVWISLELAHLKDPYALBDK 1418
Dy 1400 AEIGLHLDRSLTITIPQDPVLFSGTLRFNLDPFNHYSGDITWKTLEMANLKEFAFANEQ 1459
Qy 1419 LDHECAGGENTLVGKQOLVCLARALRKTKILVLEKATAVDDLETDLLIOSTRTOPE 1478
Dy 1460 LNTYITGSGNISGQOLVCLARALKRTRVLLIDERTAAVDSITDALLIKTIRREEPAN 1519
Qy 1479 CTYVTLIARLNTIMDYTRVIVLDKGEIOEGAPSDLLQORGL-FYSNAKQAGLV 1531
Dy 1520 ATVTIARLNTIMDYRIIVLDGKVGFEFSPANLLSNRSEFYNAKQAGLI 1573

RESULT 7
T42216
multidrug resistance-associated protein homolog MLP-1 - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T42216
R:Hirohashi, T.; Suzuki, H.; Ito, K.; Ogawa, K.; Kume, K.; Shimizu, T.; Sugiyama, Y.
Mol. Pharmacol. 53, 1068-1075, 1998
A/Title: Hepatic expression of multidrug resistance-associated protein-like proteins mai
A/Reference number: Z22081; MUID:98279126; PMID:9614210
A/Accession: T42216
```

```
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1502 <HR>
A/Cross-references: UNIPROT:O68269, UNIPARC:UPI000012557A; EMBL:AB010466; NID:93242457;
A/Experimental source: strain Sprague-Dawley; liver
C/Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 32.3%; Score 3195.5; DB 2; Length 1502;
Best Local Similarity 43.1%; Pred. No. 2,2e-208;
Matches 655; Conservative 290; Mismatches 525; Indels 51; Gaps 9;

Qy 22 WNTSNPD-----FTKCFQVTVLVWVPCFYLMACFPYFLYSLSHSDRGYIOMTPLNKTKT 75
Dy 20 MNQTEQRPVAYHLLNLCFLAAGSVVPRMTIMVGLPIYLILYHNGCCYLMSLFLKTM 79
Qy 76 ALGFLMIVCMADLFYSFWERSRGIPLAPVFLVSPITLLGTTLLATPLIQLERRKQVSS 135
Dy 80 VLGSRLILLYTFNAVPLMRHRCMPQAPPELLIPVWLMTMSPATFLIHMERKGVAS 139
Qy 136 GIMTLPLVALVCLALIRSKIMTALKEADAVDFRITPFYVYSLILQVLVSCPSDRS 195
Dy 140 GLRFGV---LLCCLVPAIDTVQOASGFRQEPHLHILATYLCLSLVAAELVLSCLVDQ 196
Qy 196 PLFSETHIDPNPCDESSASFLSRITFMWITGLIVRGYROPLESGDLSLNKEDTSEOVVP 255
Dy 197 PFRSEBCKPLNCPCEAASFPKAMFWAGSLKGRKILGPKDLSLERNSEBELVS 256
Qy 256 VLVMNKKKCAKTRKQPVKVVYSSKDPAPQESSKVDANEVEALIVKSPQKWNPSLFX 315
Dy 257 QLEEMERN-----PSELGKHGHSQM--GTPEYAFLL--QPERSRQGRPLLR 299
Qy 316 VLKTPGPYFLMSFFFAIHDLMMFSGPQILKLIKLVNDTKADWQGYFTYLLFYTAC 375
Dy 300 AVMWVFSTPLGLTLVIVIDARFAVAPKLSLFLFEMGDLSESAWGMVLAVMFLSAC 359
Qy 376 LQTLVHGYHICFVSGMRIKTAVIGAVYKALVITNARSKSVGEIVNLSVADARFM 435
Dy 360 LQTLFEOQYMRVKVLOMRIRLTATGLVYKVLVLSGSSKSSAGAVNVLVSDVQRLV 419
Qy 436 DLATYIMINISAPLOVILALYLLMLNGPSVLGVAVMVLPVNAVAMKTKTYQVAHM 495
Dy 420 ESIHLNGLWLLFLMIIVCVLQGLGPRALTAVANFLSLPLNFIITKRSHQDEQM 479
Qy 496 KSKDNRIKLMETILINGIKVLKLYAMELAFQKVLAIPOBELKVLKKSAYISAVGTFTWC 555
Dy 480 ROKASRARLITSSMLRTYRTIKSHGWECAFLERLHIGOEIGALKTSAFLPVSALVFOV 539
Qy 556 TPFLVALCTPAVYVTTIDENNIIDQTAFFVSIALFNILRPLNLIIPWYISSIVQASVLKR 615
Dy 540 STFLVALVFAVHTLVAMEDNAMDEKAFVTLVLSLNKQALFSPVHCLVQARVSPDR 599
Qy 616 LRIFLSHEBELPDSIERRPVKGDTNISTVRNATPTWASDPEPTLNGITFSPREGALVA 675
Dy 600 LAAPFLCLEEVDPMGVLPSPR-CSSKORISIHNGTFVMSQESPCHGIMLVPOGCLLA 658
Qy 676 VVGQVGGKSSLLSALLAEMDKYBGHVIAIKGVAVYVPOQAMIONDSIRENLTREGQLEBP 735
Dy 659 VVGPGVAGKSSLLSALLGELLKVEGVSIGSVAUYVQBAVMVQNTSVENAVCFROEIDL 718
Qy 736 YNSVIOACALLPDLLEITPSGDRTEIGEKVNLGGCKQKQVSLARAVYNSADITLPDPL 795
Dy 719 WLQEVLEACALGSSVVSFPAGVHTPVGEQGMNLSGGCKQKRLSARAVYRRAPVYLLMDPL 778
Qy 796 SAVDAHVKHIFENVIGPKMKLNKTRILVTHSMSTYLPQVDVIIWVGSKIEMSGSYOEL 855
Dy 779 AALDAHVSQGVFKOVIGPSGLQGTTRITLVTHLVLPQADQILVLANGTIAEKVSQDL 838
Qy 856 LAADGAFAELRTYASTEQODABENGV-----GVSGGCKAKQKQENMLVTDSAGKQ 909
Dy 839 LHRNGALVGLLDGARQPAQGEBAHAHAATSDDLGGSGGCTPTTRPE----- 886
Qy 910 LQROLSSSSSYSDISRHNNSTAELOKAKEKEETWKLMEADRAQTCQVTLVYWDYMK 969
```

[illegible]

	C,Keywords: Atp, nucleotide binding, P-loop, transmembrane protein, yeast vacuole
	F.1287-308/Domain: transmembrane #status predicted <TM1>
	F.1345-366/Domain: transmembrane #status predicted <TM2>
	F.1421-442/Domain: transmembrane #status predicted <TM3>
	F.1446-467/Domain: transmembrane #status predicted <TM4>
	F.1534-555/Domain: transmembrane #status predicted <TM5>
	F.1558-580/Domain: transmembrane #status predicted <TM6>
	F.1663-670/Region: ATP-binding cassette homology <ABC1>
	F.1951-972/Domain: nucleotide-binding motif A (P-loop)
	F.1995-1016/Domain: transmembrane #status predicted <TM8>
	F.1068-1088/Domain: transmembrane #status predicted <TM9>
	F.1092-1113/Domain: transmembrane #status predicted <TM10>
	F.1119-1200/Domain: transmembrane #status predicted <TM11>
	F.11208-1229/Domain: transmembrane #status predicted <TM12>
	F.1289-1483/Domain: ATP-binding cassette homology <ABC2>
	F.1306-1313/Region: nucleotide-binding motif A (P-loop)
Query Match	28.7%; Score 2838; DB 1; Length 1515;
Best Local Similarity	40.3%; Pred.No. 4,5e-184;
Matches 632;	Conservative 295; Mismatches 495; Indels 146; Gaps 27;
QY	28 DFTKCFQNTVAVWVDCFYIACFPYF-----LYLSHD-----RGYIOMTPLYKTKT 75
DB	27 DFTQCFIDGVIL-----NLSAIFMTTFGRDVLNCKKSHSIXKRRNIIYSRMAYLL 81
QY	76 ALGFLWITVCWADLFYSFWERSRGIPLAPVFLVSPTLGITLLATFLIQLERRKQVSS 135
DB	82 ELAF-----VSLASINISKEAEN--FTVSGYASTMLSLFVALALHWIEYDR--SVAN 132
QY	136 GIMLTFWLVAVCALAIRSKIMTKEDAOVDLRDITFYVFS-----LLLIQVLVSC 190
DB	133 TVLTFYWMLEFFEGNPAKILNIIIRHTYEG-----TWISGQTGFILTLFVGVITC 180
QY	191 FS-----DRSPL-----FSETHDNPDCPESSASFLSRITFWMTIGLIVRGYRPLEG 238
DB	191 ASILLEALPKKPLPHQHIIHOTLRKRKPNIPDSANISRTIFSMWSGLMKTGYEKLIV 240
QY	239 SDLWSLNKEDTISEQVVPVLVKNWKKCAKTRKQPVKYVSSKDPAPQKSSKVDANEVE 298
DB	241 ADLYLTPNPFSSSEELSQLEKKNEMEL-----KQKS----- 271
QY	299 ALIVKSPQKEMNPISFLKVLKTYTGGYFLMSFFPKAIHDLMFSGPOILKLIKLVND--- 355
DB	272 -----NPSLSMAICRTFGSGKMLLAEPFKAIHDVLAFTQDPLRLILIKFVLYNS 320
QY	356 -----TKAPDMQGYFTVYLFLPVAACLOTLVLHGVFHLCPFSQMRKTA 398
DB	321 ERQDDHSLIQGEENNHPOKLPVIRGGLLAFAPFLVGFOTSVLHQFLVAVNTGMATKSA 380
QY	399 VIGAVYRKALVITNSARKSSTVGEIVINLMSYDAORFMDLATYINMIWASPLQVITALL 458
DB	381 LTALLYQSLVLSNBSAGSSGSDIVNLSYVQKLODITQWNLNIMWSBPFOIICLYSL 440
QY	459 WLNLGPSVLGAVAVNVLVPVAVNAMKTYTYQVAHMSKDNRIKLMELNGIKVLKLY 518
DB	441 YKLINSMWVGVIILIVIMPLINSFLMRLOKLOQSKQMKXDESTRVYISIIANNIKSLKY 500
QY	519 AMELAFKQKVLAIR-QEELXVLKKSAYLSAVGTFTWCPPLVALCTEPAVYVYITDENNI 577
DB	501 AAEKQYREGLVEVRNKKELNLTQKGYCWAIVTSFGQNIYVFLVSCCTPAFV-YTDDRAL 559
QY	578 DAQTAFAVSLATFNILRFPNLILPMVISSIVQASVSLKRLRIFLSHELEPDSIERPVXD 637
DB	560 TTDLVFPALTLFNLISFLMITIPWILNISTEASVSGIRLFTPTNNELOPDSQRLPKYK 619
QY	638 GGGTNSITV-RNATFTWASD--PPTLNGITFISIEPGALVAVAGVCGCKSSLSALLAE 694
DB	620 NIGDVAINIGDATTFLMQCKPEYKALKNINFOAKKGNLTCTIVGKVGSGKTAIISCMIGD 679
QY	695 MDKVBGHVAIKSSVAVPQOAMITONDSLRENLILFGQLEBPYKRSYIQACALLPDLIELP 754
DB	680 IFRVVGFAIVHSSVAVVSGVPMIMGIVENILFGIRYVAEPEKTIKCAKALLIDALITLM 739

QY 755 SGRTEIGKGVNLSGGOKORSLARAVYSNADLYLPDDPLSAVDHVGKHIFFENVIGPK 814
 DB 740 DGDITVIGEKISLSGGOKARLSLARAVYADLYLDDPLAAVDHVARHLIETHVGN 799
 QY 815 GMLNKRILVTHSMYLPPOVDVIVMSGKISEMSYOEELLA-RDGAFAEFLRTVASTE 873
 DB 800 GLHFKTKVLATNKVSLNADSIALLDNGEITTOGVTDEITKXADSLMMLANNYK-- 857
 QY 874 QEOAENGVTGVSQPKAKOMENGLVYDSAGKOLQROUSSSSSYSGDISRHNSIAE 933
 DB 858 -----KNNKSNERNEDSSESSRESSIPV--EGELEJOQLKNDLDFGNSDAISLRADAT 911
 QY 934 LQKAKEEETWKMEADKQOTGOVKLSVYWDYKKA-----IGFISPLSIFLMCHNS 988
 DB 912 LGSIDFGDDE--NIAKEHEQGVKNMITYLEKAKCNPKSCVFIIFIVSMFL----- 964
 QY 989 ALASNYLSMTWDDPIVNGTOEHTKVLISVYALGISOGLAVFGYSMAVSI--GGILASRC 1047
 DB 965 SVMGVNVLLKHMSEVNSRYGSPNAPALATYFALGIGSALATLQITVLWVFCIHSKY 1024
 QY 1048 LAYDLHSIIRSPSPFERTPSGMLVNFPSKELDTVDSMTPEVIKMGSLFENVIGACIV 1107
 DB 1025 LHNIMTNSVLRAPWTFEETPIGRIILNRFSDIYKVDALIGRTPSQEFVNAVKTFPTTV 1084
 QY 1108 ILATPIAIIIPELGIYFVFOVYVASSRQLKLESVSRSPPYSHFNETLGVSTIRA 1167
 DB 1085 ICATTWOFITIIPLSVFYIYQOYVLTSTRELRDLSTIRSPYSHFOETLGLATVVG 1144
 QY 1168 FEOBERFIHOSDLKVDENOKAYPSIVANRMLAVRELCVNGCIYLFALPAV--ISRHS 1225
 DB 1145 YSQOKRSHINOCRIDNNMAFYPISANRMLARLLIGSIITLGAATLSVPRKCGTL 1204
 QY 1226 SAGLVGLSVSYSLQVTTYLWMLVMSSEMETNIVAVRLKXSETEKAPMOIQETAPS 1285
 DB 1205 TAGVWGSLSTYALQITQLNMIYAMTVEVETIVSVRIKEXADLKSEAPLIVGHRPK 1264
 QY 1286 SMPQGVVERFRNYCLARRELDLPYLRIHINTINGEKVGIYGRGAKSSITLGLFINE 1345
 DB 1265 EWSQGGIKFNNSSTRYRPELDVLKIHINHIKENEKVGIVGRGAKSSITLGLFINE 1324
 QY 1346 SAGEEIIIDGINIAKIGLHDLRFKTIIPDVPVFGSLRMNLDPFQYSGDEEVTSL 1405
 DB 1325 ASEGNIIVINDIATNEIGLYDLRHLSTIPDSQVFEETVARENIDPINOYIDEATWRLEL 1384
 QY 1406 AHLKDPVSALP-DKLDEHCAEGENLSVGOQOLVCLARALLRKTKIIVLDEATAVDLET 1464
 DB 1385 SHLKEHLSMSNDGLDAQLTEGGNLSVGOQOLCLARALLRKTKIIVLDEATAVDLET 1444
 QY 1465 DDLIQSTIRTOFEDCTVLTAAHRLNTIMDTTRYVYLDKGEIOEYGAASDL--QORGLFYS 1523
 DB 1445 DKVQETIRTAFODRITLTIAHRLNTIMDSRIIVLDNGKVAEFPDGLSDNKSLEY 1504
 QY 1524 MAKDAGLV 1531
 DB 1505 LCMERAGLV 1512

C:Genetics:
 A:Gene: CESP.E03G2.2
 A:Map position: X
 A:Introns: 73/3; 113/3; 176/1; 216/1; 258/2; 384/3; 470/3; 549/1; 644/3; 757/3; 764/2;
 C:Superfamily: human multiting resistance protein cMOAT2; ATP-binding cassette homology

Query Match 27.6%; Score 2732; DB 2; Length 1398;
 Best Local Similarity 41.1%; Pred. No. 6.5e-177;
 Matches 591; Conservative 287; Mismatches 472; Indels 88; Gaps 19;

QY 115 ITTLATFLQLERKRVOSGIMLTFWALVALCALILRSKIMTALKED--AQVDLFRD 172
 DB 27 LALLIATNEV---RRAGIHSSGFLPCIMMLFAVAADVEFQMMTGSQPELVADIFRY 83
 QY 173 ITTYVYSLLIQVLVSCFSDRSLPSETIHDN-----PCPESSASFSLRTWITGL 227
 DB 84 VAYLYFPLVAVAEFVLFHVSDFPFM-----PRGYONLCPREENANFISROLIMFTQI 136
 QY 228 IYRGYRQPLSGDLSLNKEDTSEQVYVLYNMKKECAKTRKQPVK-----YSS 279
 DB 137 ISLGEKTLVADDVFEMDSQMDQBYLKAARKTEMLKOTERAKQVRLDKREARATGSE 196
 QY 280 KDAOPKSSKVDAN--EEVEALIVKSPQKEMNPSLFVLYKTYGTPYFLMSEFFRAIIDL 338
 DB 197 KAPLGTFFNNYGAVNDLDDKDRIVQ-----PSVITLWQIMKWEILGGSFIKFLSDL 249
 QY 339 MFGPQILKLLIFVNDTKAPDMQGYFYVLLVYVACLOTLVHQYHICFVSGMRKTA 398
 DB 250 QFANPFLNLYLIFIEFPNAPLNLINGIGLAEPLAQIKSLFMNTYFIAMTRVAKIOTM 309
 QY 399 VIGAVRKALVINRSKSTGEIYNLMSVDQRFMDLATYINMTMSAPLOVLYL 458
 DB 310 LSCAVYKSLSLNTARREFTVEMNILSIDRFMTIPQIQOYSSPFOIITCVLL 369
 QY 459 WNLGSPVLAVGAVVWMLVFNVAVMAMKTYOVAAHMSKDNKIKLMBELINGIKVLY 518
 DB 370 SOTIGAVVAGIYVMSIYVINICVMITRMQLMKYDEIRLNEVTLNGIKVLY 429
 QY 519 AMELAFKDYLAIRQBELKYLKSAVLSAVGTFTWCTPELVALCTPAVYVITDENNILD 578
 DB 430 AMETAMEETIERVDEKELMKIKQSALLKTFADCLANGCAPVFAVLSPTFEVLLIDPKVLT 489
 QY 579 AQAPFSLMFLNLRPMLNPLRVISIVQASISLRLRFLSHEBELPISIRRPYKDG 638
 DB 490 PNIAFVSLFNLRLRPLMMAELVAQTVQVVSNNRVRFLEKEVDNADIKK--IRGE 548
 QY 639 GGTNSITVRNATPTMARSDPPTLNGITPISIPEGALVAVVGVGCGKSSLSALLAEMDKV 698
 DB 549 LYNTVYEHISGSPANSAERILSDIEFLAGSELYTVVGVSGSSLSLLALGEMEKV 608
 QY 699 EGHVAIKGSVAVYVQQAQWIMQNSLRENILFGQLBEPPYRSVIOACALLPDLILPSGDR 758
 DB 609 CGYGVGRGSVAYLSQPFMLINQSLKKQVLMQADLDVLYKYVIESCALKEDLKQLPDGD 668
 QY 759 TEIGERGVNLSGGOKORVSLARAVYSNADLYLPDDPLSAVDHVGKHIFFENVIGPK 818
 DB 669 TEIGERGVNLSGGOKARIALARAVYQKDYVFLDDPLSAVDHVGKHIFFENVIGPK 728
 QY 819 NKTRILVTHSMYLPPOVDVIVMSGKISEMSYOEELLARDGAFAEFLRTVASTEODDA 878
 DB 729 HTTRILVTHSTSLGSGKIIYWK-----ETREBSG 761
 QY 879 EENGVTGVSQPKAKOMENGLVYDSAGKOLQROUSSSSSYSGDISRHNSIAE 938
 DB 762 BEEG-----GENSDIIPGSIASGSRMSRLSR--LSK-----ISRKKSISYK 804
 QY 939 AKKEETWKMEADKQOTGOVKLSVYWDYKKAIGLFSFISIFLMCHNS--ALASNYLS 997
 DB 805 -KRPDA--LITKEBAIGRVNPGVYLLYFPAMGIVTVYVYALAVLVANSFALGRSLMT 861
 QY 998 LMTDPIVNGTOEHTKV--RLSVYALGISOGLAVFGYSMAVSI--GGILASRC 1055
 DB 862 AMSDANIDINHPPTMVGARLGVYAGFTIEVIFLPSVLVLLIGVAVSKNHLKRLN 921

Db 1088 DSIRSPPVYAOFGELNLSLITIRAYKAYDRMADINGRSMNNIRFTLVNMGANRMLGIRL 1147
Qy 1204 ECVGNICIVLFAALFAVI-----SRHLSAGLVGLSVSYSLQVTTYLNMVLRMSSEMETNI 1258
Db 1148 ETLGLGLMIMLTAFAVWNGRAENQAFASWTWGLLSVALNITSLTGLVLRSLAENSL 1207
Qy 1259 VAVRLKRYSETEKAPMOIQETAPSSMPVOYGRFERNYCLIRREDDFVLRHINVTIN 1318
Db 1208 NAVEVGVYIEIRPEAPRVIEENRPPCPMPSSGSIKFEDEVLRYPOLRPVLRHVSFEIH 1267
Qy 1319 GGEKGVIGRTGAGSSLTGLFRINESAGEIIIDGINIAKIGHADRKFRTIIPOPV 1378
Db 1268 PTDKVGIVGTGAGSSLLNMLFRIVEYKGRILIDCCVCGKFGMLDKRGILGIIPOSPV 1327
Qy 1379 LPSSGLRNMULDPFQSYDEBVTSLSLAHLKDFVSLPDKLDHECAGEGEMLSVGORQV 1438
Db 1328 LFGSVTRNLDPPFGHNNADLMESLEPAHLKOTIRNRLGDAEVSAGEENSVGQRLL 1387
Qy 1439 CLARALKRKILVLDETAVADETDLIOSTIRTOGEDCTVLTIRRLNTIMDYTVI 1498
Db 1388 SLSRALLRRSKILVLDETAVAADVTDALIOKTIEEFKSCMTLIIAHLNTIIDCKIL 1447
Qy 1499 VLDKGEIOEXGAPSDLLQORGLFVS-MAKDAG 1529
Db 1448 VLDSGRVQEFSSPENLLSNEGSSFKMWQSTG 1479

RESULT 12

glutathione S-conjugate transporting ATPase (AtMPPI) - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: D86428
R/Theologian: A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasey, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nucleotide 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D86428
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1622 <STO>
A:Cross-references: UNIPROT:Q9C8G9; UNIPARC:UPI000009CF47; GB:AE005172; NID:G11055814, F
C/Genetics:
A:Map position: 1
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 23.8%; Score 2356.5; DB 2; Length 1622;
Best Local Similarity 35.6%; Pred. No. 2.9e-151;
Matches 569; Conservative 274; Mismatches 564; Indels 193; Gaps 31;

Qy 47 WACPFVFLVSRDRGICWMTPLNKTALG-----FLMTVCADLFYSFW 94
Db 8 WYCKEVENGVMTKTVDAVGAATPCADISFVLGISHLVLLILCTRLMLITTDHKKVDFC 67
Qy 95 ERSR--GIFLA-----PVFLVPTLLGITT-----LATFLIQERERKQVS 134
Db 68 LRSKFSFSLMLLAAYATBPLFRL--VMRISVLDLDGAGFPPEAFMLVLE--AFAMG 122
Qy 135 SGIMLTF-----WVVALVCAALIRSKIMTALKEDAQVDLFRDITVYVPSLL 183
Db 123 SALVMTVETKTYIHLEPMYVFAVYVAVGDMVLNLVSVK--EYVGSFKLYLISEVA 181
Qy 184 IQLV-----LSCFSDRSPLFSETH-----PNCPGESASPLSLTFMW 223
Db 182 VQVAFGTLTFVYFPMULDPYGYTPVGTENSEDEYEEELPGENICERHANLFDISIFSW 241

Qy 224 ITGLIVGYPOLGSDLSLNKEDTSEQVVPVLYKMKKCACTRKPQVNVYSSKDA 283
Db 242 LNPMLTGSKRPLTEKQVWMLDITWDTETLRSPFKSMDEKLEK----- 285
Qy 284 QPKSSKVDANEVEALIVKSPKEMNPSPLEKVIYKTFGAPFLMSFFFKALHIDLMFSGP 343
Db 286 -PK-----PWLRALNLSIGKFFMWGGEFKTGDCSQFVGP 320
Qy 344 QILKLLIKFVNDTPAPMVGYYTVLLFVTAQLQTVLHQYFHI CFVSGMR IKTAVIGAV 403
Db 321 LLNLMLKSM-QLMBPAMIGIYVAISIFVGVLVGLCAQYFQVMRVGYRLRSALIAV 379
Qy 404 YRKALVITNSARKSTGYEIVNLSVDAQRMDLATIYINMTWSAPLOVITLALMLNLG 463
Db 380 FRKSLRLTNEGRKRFQTKITNMTTDAESLQQICQSLHTWSAPFRITVALVLLYQGLG 439
Qy 464 PSYLAGVAVWMLVYVNAVMMKTKTQVAMKSKDNRIKMLNELINGIKYLKYAMELA 523
Db 440 VASITIGALFLVLMPIQTVIISKQKLTKEGLQRTDKRIGLMBEVLAMDVCKIAMENS 499
Qy 524 FKDKVLAIROBELKVKLKKSAVLSAVGTFVWCTPELVALCTPAAYVITIDENNILDAQTAF 583
Db 500 FQSKVQVTRDELQMPFRAQLLSAFNMFILNSIVLVTVVSGVFSLLGSD--LTPARAF 557
Qy 584 VSLAFNLRPPLNPLMNVISIVQASVSLKRLAIFLSHEE--LEPDSIERPKVDGGCT 641
Db 558 TSLSLFVLRPPLMPLPVIITQMVNANVSLNRLEEVSLTEERVLIPND---PIEP--GQ 611
Qy 642 NSITVRNATFTW--ARSDPPTLNGITFSPREGALVAVVQVCGKSSLSALLAEMD-KVE 699
Db 612 PAISRNGYFMSDADPFLSNINLDPGLSLAVAVSGTSEGKTSLSANLGLIPARSD 671
Qy 700 GHVAIKGSVAVPOQAMIONDSURENLIIFGCOLEPPYRSVIOACALPDLIELPSSGDR 759
Db 672 ATVTLRGSVAVVPQSWIFNATVRDNLIFGAPRPOEKERYIDVTALQHDLELLPGDGLT 731
Qy 760 EIEGKVNLSCGQQRVSLSAAVYNSADVIYFDPLSADVDAHVKHITENVYIGPGMLKN 819
Db 732 EIEGKVNISGQQRVSMAVAVYNSVNCIIDLPLSLDHAHVQGVQVEKCI--KREDGQ 789
Qy 820 KTRILVHSMYLVQVNDVIIMSGSKISEMSYOELARDGAPAEFLRTVASTEOQDAE 879
Db 790 TTRVLTVNLQHLPLSQVDKILVHEGTVAEBGTIELCHSGPLFQRLMNAKGVEDY--SE 847
Qy 880 ENGVTVSGPKEKAKOMENGMVYDSAGQOLQRLSSSSSYSGDISRHNNSTAELOK--- 936
Db 848 ENGEAEVD--QTSXKPVVNG-----NANNLQXOGI 875
Qy 937 -AEAKKEETWLMEDAKQGTQVNLVYWDYMKAI-GLFISFLSIFLMCHNVSLASNY 994
Db 876 ETKNSKEGNSVLYRKEERETGVSWKVLERYQNALGAMVMMVLVICYLTQVFEFVSSST 935
Qy 995 WLSLMTD--PIVNGTOHTKRLSLVYGALGISQGINAFVGMVNSIGGIIASRCLHDL 1052
Db 936 WLSEWTDGCTKTHGRLPYN---IVALLSFGQVSVTLINSYWLIMKSLTAACKWHDAM 991
Qy 1053 LHSILRSPMSFEERTPSGNLVNRFSEKEDYDMSIPVIYKPMGSLFVNIAGCIYLLAT 1112
Db 992 LGSITRAWVFFQNPGLRIINRPAKMGDIDRYAVAVNMFMGSIADLLGVLIGIVS 1051
Qy 1113 PIALIIPPLGLIYFFQRFYVASSRQLKRLSEVSRSFVYSHNETLLGVSYIRAEBOE 1172
Db 1112 RMAELNGSMNNNRFTLVNMAANRMGLIRLEVLGGLVWMLTASLAVWONGKAQAQOAYA 1171
Qy 1228 GLVGSVGSYLOVTTYLNMVLRMSSEMETNIIVAVRLKEYSETEKAPMOIQETAPSSM 1287
Db 1172 STMGLLSYALSTISSLTAVLRSLAENSLNSVRGVNVIETIPEALVLIENRPPGWM 1231
Qy 1288 POYGVREFRNNTCLARYREDDFVLRHINVTINGEKVGIVRTGAGKSSLTGLFRINESA 1347

Db 1232 PSSGSIKEDVLRKRPFLPVHGVSPFLSPMDKVGIVGRTGAGKSLNALPRIVLE 1291
 Qy 1348 EGEIITIDGINAKIGHDLRFKRTITIPDDPVLFSSGLMNTDPPSOQSDDEEWTSLIAH 1407
 Db 1292 KGRILIDECIDIGRGLMDLRKVLGIIPQAPVLFSGTVAFNIDPPSEHNDALMWSLEIAH 1351
 Qy 1408 LKDFVSLPDLDBHCAGENGLSVGRQLVCIARALRKTIIVLDEATAVLETDL 1467
 Db 1352 LKDTIRNPLGLDEVTAGENFVGQRQLSLAALRKSKILVLDATAVDRIVL 1411
 Qy 1468 IQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLKDGEIQEYCAPSDLLQORGLFSPMAKD 1527
 Db 1412 IQKTRBEFKSCCTMLIAHRLNTIIDCKVLVDSGKQEFSSPENLINSBESSRSK--- 1468
 Qy 1528 AGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAEPRTRE 1567
 Db 1469 -----MVOSTGTANAEYLRSLTLENK-----RTRE 1493

RESULT 13

E86428
 Probable ABC transporter (imported) - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: E86428
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chun, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewart, K.;
 Ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, Y.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luroso, J.S.; Maiti, R.; Marzilli,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: E86428
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-1495 <STO>
 A:Cross-references: UNIPROT:Q9C8H0; UNIPARC:UPI00000482FF; GB:AE005172; NID:g11055818; F
 C:Genetics:
 A:Map position: 1
 C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 23.5%; Score 2326; DB 2; Length 1495;
 Best Local Similarity 34.3%; Pred. No. 3e-149;
 Matches 541; Conservative 314; Mismatches 566; Indels 154; Gaps 28;
 Qy 2 ALRGFCG--ADGSDPLMDNVTNNTSNDFTKCFONTVLVWP-----CFY-LMACFP 51
 Db 5 ALNNYCPVADGF-----MEKAVDGAFCAGAYTPCAIDSLMLVSHFVLGLCFYRIWIF- 58
 Qy 52 FYFYLISBHDGTYQMTPIKTKALGFLMTIVCMAD-----LFYSWESRGIPLAPV 105
 Db 59 -----HNTAQIYVRKKYKNCVLG-LACYCVEEVLRLVNGISLFMDDETDPPPF 110
 Qy 106 FLVSPFLGLITTLATFELIQERRGVQSSGIMLTFLMLVALVCLALIRSKIMALKEDA 165
 Db 111 EVASLWAEAFWFMVLMVIGLETQYKVEFMYVRFGLVYLVAADVLDLVLKXIN 170
 Qy 166 QVDFLRDIT--FVYVSLLIQVLSCFSDRSLPSETTHDP-----NCP 209
 Db 171 RTALYLRISRCGALFGILLIYIPBL--DPYEGYHVNNEPLDNVEYDALRGHEHICP 228
 Qy 210 ESSAFPSRITTFMWTGLIVAGYRQPLSGSDLMGLNKEDTSEGVVPLVKNMKKECAKTR 269
 Db 229 ERHAFISRIYFGITPLMQIGRKPLTEKDWQMDQETELIKAFQRCWTESSR-- 286
 Qy 270 KQPVKVVYSSKDPQPKSSKYDANEEVALIVKSPKWNPSLFKVLKTFGYPFLMSF 329
 Db 287 -----PK-----PMLRLALNNSLGRFNLAG 307

Qy 330 FFKAHDLMMSSGPOILKLIKFNVDTKAPDMOGYFTVLLFYTACIQTLVLAHOYFHICF 389
 Db 308 IFKIGNDLSQFVGVIILSHLRSMQEGD-PAMVGYVAFTIFVGTGLVCEAQFQNW 366
 Qy 390 VSGMRITAVIAGVYRRALVTNSARKSFTGEIVINLMSVDAQRFMDLATYINNIWAPL 449
 Db 367 RVGRRLASTVLAALFHKSLRTHARKNFPASGKTYNITTDANLQOISQDLHGMSAPF 426
 Qy 450 QVIALYILMLNIGPSVLAGVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 509
 Db 427 RIIVSMILLYQQLGVASLFGSLIFLLIPLIOTLIISKMRKUTKEGLQWTDKRVGTNEIL 486
 Qy 510 NGIRVILKLYAMELAFKQVLAIRQBEIKVLKKSAYLSAVGFTVTCIPFLVACTFPAAVY 569
 Db 487 SSMOTVVCYAMEKSFESRIQIRNEELSWERKQDLAFNSFILNSIPVVTVVVSFGVAV 546
 Qy 570 TIENNNILDAQTAQVSLALFNILFEPNIPILPMVSSIYQASVSLKRL-RIFLSHEBELPD 628
 Db 547 ILGSD--LTPARATSLSPAVLRFLPMLPNLISQVYVAVNSLQRIEELLSERILAQ 604
 Qy 629 SIERRPVKDGGINSTIVRNATFTW-ARSDPTLNGITFSIPGALVAVVQVCGKSSL 687
 Db 605 NPLQF-----GTPAISIKNGYFSDGKTKTPTSDINLIPVGTIVAVIGTGEGKSTL 659
 Qy 688 LSALLAEMDKVE-CHVAIKGSVAVYPOQAMTQNDLSLENNILFGCOLPEPPTVYIQCAL 746
 Db 660 ISAMIGELSHAETTSVIRGSVAVYPOVSWIFNATVRENILFGSDFESERYARIDAVAL 719
 Qy 747 LPDEILIPSGDRTIEGKGVNLGGQKORVSLAAVYNSNDIYFDDPLSLAVDAHVGHII 806
 Db 720 QHDDLDLPGRLDLEIGRGVNIISGQKORVSAARVSNISVYIFDDPLSLADHVAHQV 779
 Qy 807 FENVIGKGMKNKRLIVLTHSMYLRPOVYVIVMSGKISEMOSYOEALARDCAFAPFL 866
 Db 780 FDSGM--KDELRGKTRVLVNLQHLFLMLKXIIIVSSGMIKEBTFFELSLSGILFKML 837
 Qy 867 RTASTQEBQDAEENGVTVGSGPEKAKQEMENGLVTDAGKQLQROLSSSSYSGDISR 926
 Db 838 ENAKMDATQEVNTN-----DENILKLGPTVTVDS--ERNIGSTR----- 876
 Qy 927 HHNSTAELOAABAKSETWLMADKQOTGVLSYVWDYWKAI-GHIFISLIFLPMCN 985
 Db 877 -----QGRKRS-VLTKOERETGIISNNVLRKYEAAGSLAVMMILACYAT 924
 Qy 986 HVSALASNYMLSTWDDPIVNGTQEHK-----VRLSVGALGISQGIAGVGSMAVSIG 1040
 Db 925 EVLRVSSSTWLSIWTD-----QSTSKNYSPPGYIVVYLLIGFQVAVFTNSFWLITS 977
 Qy 1041 GILASRLHVDLHSLIRSPMSFFERTPSGNLVNRSKELDPTVDSMTPEVIKPMFSLFN 1100
 Db 978 SLHAARLHAPMLSSIRAPMLFPHTPTRVIRFESKDIGDDRANVANNMFMNDLWQ 1037
 Qy 1101 VIGACIYILATPIAIIIPGLIYFVQGRFYVASSROKLRESVSRSPPYSHNETL 1160
 Db 1038 ILSTPALIGVSTISLWALPMLLILFYAALVYQSTREVRRLDSVTRSPITYAOFGLIN 1097
 Qy 1161 GVSIVTAFEEOEPIHOSDLKVDENQAYPSIYANMLVRLCEVCNCLVLPALPAVI 1220
 Db 1098 GLSSIRAYKAYDRNAKINGSMONNIRFTLANISSNNMLIRLETLOGWIMWLTRAVL 1157
 Qy 1221 SRHSLG-----AGLVGLSVSYLQVTTYLMLVMSSEMETNIVAVRLKEYSETEKAP 1275
 Db 1158 QNGNTNNQAGFASMTGMLLTSTYTNISLISLGGVLRQARARANSINSVERVGVIYDLPSEAT 1217
 Qy 1276 MOIQTAPSPSPVGVGEFRANCLRYRELDLFLRIINTVINGEKGAGVIGRTGAGKSS 1335
 Db 1218 DLIENRFPVCGMPSGSGISKEVDHRLRPPGLPPLHGLTFFVPSSEKYGAVGRGAGKSS 1277
 Qy 1336 LTLGLFRINESABEIIIDGINIAKIGHDLRFKRTITIPDDPVLFSSGLMNTDPPSOQS 1395
 Db 1278 MLNALPRIVEVKGRIIMIDCDVAKFGLTVRRLVLSIIPSPVLFGSTVAFNIDPPSEHN 1337

[illegible]

Db 530 ---KSEKRGQXKEKNE-- NVQLGRYKKSYYQLYIKMGIFNSAPLIFLAHFTVLMIR 584
QY NYWLSLMTDD-----PIVNGTQE-----HTKRLSVYGAIGISOGIAVFGYSM 1035
Db 585 SLMLDWSNENNAIKKATLLSVLYLNSSTSSVDGVPVSVETRLIYVAGFG--GLEMLLAL 641
QY 1036 A---VSIGGIIASRCIHDVLLHSILRSPMSFFERTPSGNLVNRSKELDTVDSMPEVIX 1092
Db 642 APTVTLIGSLRASVGHSLPIHLLVAPISFPTTPGRIINRLSRDLVDIDK-LQDNIR 700
QY 1093 MEMGSLFNVIAGACIVILLATPIAIIIPPLGIYFFQRFVVASRQLKRLSESRSPPY 1152
Db 701 MCTGTLINACMLIVLISITPIFLVCAAPILILIYFPMIYIIPISROLKRLSANSRPL 760
QY 1153 SHNETLLGVSVIRAFEEQERFTHQSDLKVDENQKAYYPSIVANRMLAVRLECVGNCTVL 1212
Db 761 STIASIHGASSIRAPDKTERTTALTSTNDKFAQCRYLSHMSNRMLATRLLELGNCTVL 820
QY 1213 PAALPAVIR--HSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETNVAALRLKEYSBT 1270
Db 821 FASLSATSTSKYFGLPPGAGLSVSYALTTEVTNLCVRSVSELSNVSVERVNEYQL 880
QY 1271 EKEAPMOIETAP-PSSWEPQVGRVFRFNRYCLARYEDDLFVLRIHNTVINGEKVAVIGRT 1329
Db 881 EEAPAPRIEKSLENEKMPVYKGIELDGFMSRYKRLPLVKNLIDKIEGGERIGVIGRT 940
QY 1330 GAGKSSLYLGLFRINESAGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLPFGSLPMND 1389
Db 941 GSGKSSLYLALYRMIGESGTTIKIDVEIDTIGLHQRSKLIIIPQEPVSVFSGTLRFMND 1000
QY 1390 PPSQVSEDEWVMSLELAHKDPFSALPDKLDHECAGEGNSVSGOROLVCLARALLRKTK 1449
Db 1001 PENQVSDDDIIMWCLEICQLKQFAQEDDKTLIDRYIABGSKNMSVBRQLCLCRALLRGAR 1060
QY 1450 IYVLGEATAVADLETDDLQSTIRTFQEDCTVLTIAIRLNTIMDYTRVIVLDKGEIOEXG 1509
Db 1061 IYILDEATASVDTVTDGIVQRAIRQHPPOSTTISIAHRLDTIVSDRLVLDAGVAERD 1120
QY 1510 APSDLLQQRGLFYS 1523
Db 1121 TFSNLLNPDLSYS 1134
RESULT 15
S64757
probable membrane protein YIL015w - Yeast (Saccharomyces cerevisiae)
N.Alternate names: hypothetical protein A255; hypothetical protein Lu133
C.Species: Saccharomyces cerevisiae
C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C.Accession: S64757; S64763; PS0041; S69391; S70560
R.Mihsa, T.; Zimmermann, F.K.
submitted to the Protein Sequence Database, May 1996
A.Reference number: S64743
A.Accession: S64757
A.Molecule type: DNA
A.Residues: 1-1559 <MIO>
A.Cross-references: UNIPROT:P14772; UNIPARC:UPI0000126A9C; EMBL:Z73120; NID:g1360184; PI
A.Experimental source: strain S288C
R.Goffeau, A.; Purnelle, B.
submitted to the Protein Sequence Database, May 1996
A.Reference number: S64761
A.Accession: S64763
A.Molecule type: DNA
A.Residues: 1-1559 <GOF>
A.Cross-references: UNIPARC:UPI0000126A9C; EMBL:Z73120; NID:g1360184; PIDN:CAA97460.1; P
A.Experimental source: strain S288C
R.Boy-Marcotte, E.; Danak, F.; Camonis, J.; Garreau, H.; Jacquet, M.
Gene 77, 21-30, 1989
A.Title: The C-terminal part of a gene partially homologous to CDC25 gene suppresses the
A.Reference number: PS0041; MUID:89306677; PMID:2545538
A.Accession: PS0041
A.Molecule type: DNA

QY 1354 DGINIAKIGLHDLRFKTTIIPQDPVLESGSLRMNIDPFSQYSDEBWTSLAHKDFVS 1413
 Db 1363 DGLIDISDIGLFDLRSHLAIIPQDAQAFEGTVKTNLDPFNRYSDDELKRAVEQAHLKPHLE 1422
 QY 1414 AL-----PDKLDHECAEGGENLSVGOROLVCLARALLRKTILVLEA 1456
 Db 1423 KMLHSKPRGDDSNEDGNVNDILDVKINENGSNLSVGOROLLCLARALLRSKILVLEA 1482
 QY 1457 TAAVDLETDLIQSTIRTFQEDCTVLITAHRLNTIMDYTRVIVLDKGEIOEYGAPSDILQ 1516
 Db 1483 TASVDMETDKIIOPTIRREFKDRITLITAHRIDTVLDSDKIIVLDQGSVREFDPSKULS 1542
 QY 1517 QR-GLFTSMADAG 1529
 Db 1543 DKTSIFYSLCEKCG 1556

Search completed: December 15, 2005, 15:36:06
 Job time : 56.8285 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model1

Run on: December 15, 2005, 14:56:46 ; Search time 243.004 Seconds

(without alignments)
5594.767 Million cell updates/sec

Title: US-10-665-283-1

Perfect score: 9903

Sequence: 1 MALRGFCSADSDPLMDMNV.....RSVAVAKAKPKFSTSPDLS 1927

Scoring table:

BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

UniProt 05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7860	79.4	1531	MRP1_HUMAN	P33527 homo sapien
2	7769	78.5	1515	Q9UQ99_HUMAN	Q9UQ99 homo sapien
3	7695	77.7	1531	Q864R9_MACFA	Q864R9 macaca fasc
4	7689	77.6	1531	Q864S0_MACFA	Q864S0 macaca fasc
5	7444	75.2	1459	Q9UQ97_HUMAN	Q9UQ97 homo sapien
6	7419.5	74.9	1456	Q9UQ90_HUMAN	Q9UQ90 homo sapien
7	7272	73.4	1531	Q6UR05_CANFA	Q6UR05 canis famli
8	7242	73.1	1439	Q59G19_HUMAN	Q59G19 homo sapien
9	7230.5	73.0	1530	Q8HX05_BOVIN	Q8HX05 bos taurus
10	7104.5	71.7	1400	Q9UQ98_HUMAN	Q9UQ98 homo sapien
11	7002.5	70.7	1528	MRP1_MOUSE	Q35379 mus musculu
12	6932.5	70.0	1532	Q810B4_RAT	Q810B4 rattus norv
13	6932.5	70.0	1532	Q8CG09_RAT	Q8CG09 rattus norv
14	6892	69.6	1523	Q810G9_RAT	Q810G9 rattus norv
15	6301	63.6	1525	Q5F364_CHICK	Q5F364 gallus gall
16	6075	61.3	1215	Q68CP7_HUMAN	Q68CP7 homo sapien
17	4526	45.7	1522	Q5SUF4_MOUSE	Q5SUF4 mus musculu
18	4522	45.7	1519	Q80ZK8_MOUSE	Q80ZK8 mus musculu
19	4522	45.7	1522	Q56PH0_MOUSE	Q56PH0 mus musculu
20	4508.5	45.5	1523	Q59DU0_MOUSE	Q59DU0 mus musculu
21	4491.5	45.4	1527	MRP3_HUMAN	Q51343 mus musculu
22	4484.5	45.3	1533	Q59H05_HUMAN	Q59H05 homo sapien
23	4422	44.7	1514	Q96QA9_HUMAN	Q96QA9 homo sapien
24	4402	44.4	1498	Q59DK9_MOUSE	Q59DK9 mus musculu
25	4396	44.4	1522	MRP3_RAT	Q487R7 rattus norv
26	4148.5	41.9	1496	Q487R7_TETNG	Q487R7 tetraodon n
27	3885.5	39.2	1548	Q7KTC3_DROME	Q7KTC3 drosophila
28	3880.5	39.2	1548	Q7KTC1_DROME	Q7KTC1 drosophila
29	3879.5	39.2	1548	Q7KTC2_DROME	Q7KTC2 drosophila
30	3874.5	39.1	1548	Q7KTC2_DROME	Q7KTC2 drosophila
31	3860.5	39.0	1548	Q917N0_DROME	Q917N0 drosophila

ALIGNMENTS

RESULT 1	MRP1_HUMAN	STANDARD	PRT	1531 AA.
ID	MRP1_HUMAN	P33527; O14819; P78419;		
AC	P33527; O14819; P78419;			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	13-SEP-2005 (Rel. 48, Last annotation update)			
DE	Multidrug resistance-associated protein 1 (ATP-binding cassette sub-family C member 1).			
GN	Name=ABCC1; Synonyms=MRP, MRP1;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;			
OC	Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	NCBIOTIDE SEQUENCE.			
RX	MEDLINE=93088080; PubMed=1360704;			
RA	Cole S.P.C., Bhargava G., Gerlach J.H., Mackie J.E., Grant C.E.,			
RA	Almquist K.C., Stewart A.J., Kurz E.U., Duncan A.M.V., Deeley R.G.;			
RT	"Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line."			
RL	Science 258:1650-1654(1992).			
RN	[2]			
RP	SEQUENCE REVISION.			
RX	MEDLINE=93262415; PubMed=8098549;			
RA	Cole S.P.C., Deeley R.G.;			
RT	"Multidrug resistance-associated protein: sequence correction."			
RL	Science 260:879-879(1993).			
RN	[3]			
RP	NCBIOTIDE SEQUENCE.			
RX	MEDLINE=96251691; PubMed=8649356;			
RA	Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P.C.,			
RA	Deeley R.G.;			
RT	"Structure and expression of the messenger RNA encoding the murine multidrug resistance protein (MRP) an ATP-binding cassette transporter."			
RL	Mol. Pharmacol. 49:962-971(1996).			
RN	[4]			
RP	NCBIOTIDE SEQUENCE, AND ALTERNATIVE SPLICING.			
RX	MEDLINE=8008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;			
RA	Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;			
RT	"Analysis of the intron-exon organization of the human multidrug-resistance protein gene (MRP) and alternative splicing of its mRNA."			
RL	Genomics 45:368-378(1997).			
RN	[5]			
RP	NCBIOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA] OF 1131-1531.			
RX	MEDLINE=99425270; PubMed=10493829; DOI=10.1006/geno.1999.5927;			
RA	Lothman B.J., Kim U.-J., Shedd V.P., Kalush F., Brandon R.,			
RA	Fuhrmann U., Mason T., Crosby M.L., Barnstead M., Cronin L.,			
RA	DeLattre Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,			
RA	Etchler E.E., Harris P.C., Venter J.C., Adams M.D.;			
RT	"Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q."			
RL	Genomics 60:295-308(1999).			

32	3857	38.9	1549	2	Q7KTC0_DROME	Q7KTC0 drosophila
33	3845.5	38.8	1548	2	Q8T9C5_DROME	Q8T9C5 drosophila
34	3838.5	38.8	1548	2	Q7KTB9_DROME	Q7KTB9 drosophila
35	3836.5	38.7	1548	2	Q7KTB9_DROME	Q7KTB9 drosophila
36	3818.5	38.6	1548	2	Q7KTC0_DROME	Q7KTC0 drosophila
37	3813.5	38.5	1548	2	Q7KTC8_DROME	Q7KTC8 drosophila
38	3812.5	38.5	1548	2	Q7KTC4_DROME	Q7KTC4 drosophila
39	3807.5	38.4	1548	2	Q7KTC9_DROME	Q7KTC9 drosophila
40	3793.5	38.3	1548	2	Q7KTC7_DROME	Q7KTC7 drosophila
41	3790	38.3	1549	2	Q9VK56_DROME	Q9VK56 drosophila
42	3771.5	38.1	1548	2	Q7KTC5_DROME	Q7KTC5 drosophila
43	3769.5	38.1	1548	2	Q7KTC6_DROME	Q7KTC6 drosophila
44	3741	37.8	1564	2	Q8Q998_RAJER	Q8Q998 raja erinac
45	3738.5	37.8	1567	2	Q6PH26_BRARE	Q6PH26 brachydantio

[6]
RP TOPOLOGY, AND CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=97444425; PubMed=9295302; DOI=10.1074/jbc.272.38.23623;
RA Høfner D.R., Almquist K.C., Leslie E.M., Gerlach J.H., Grant C.E.,
Deeley R.G., Cole S.P.C.;
RT "Membrane topology of the multidrug resistance protein (MRP). A study
of glycosylation-site mutants reveals an extracytosolic NH2
terminus."
RL J. Biol. Chem. 272:23623-23630 (1997).
RN [7]
RP TOPOLOGY.
RX MEDLINE=97476249; PubMed=9334225; DOI=10.1074/jbc.272.42.26479;
RA Kaet C., Gros P.;
RT "Topology mapping of the amino-terminal half of multidrug resistance-
associated protein by epitope insertion and immunofluorescence."
RL J. Biol. Chem. 272:26479-26487 (1997).
RN [8]
RP TOPOLOGY.
RX MEDLINE=96153110; PubMed=9485377; DOI=10.1021/bi972332v;
RA Kaet C., Gros P.;
RT "Epitope insertion favors a six transmembrane domain model for the
carboxy-terminal portion of the multidrug resistance-associated
protein."
RL Biochemistry 37:2305-2313 (1998).
RN [9]
RP MUTAGENESIS OF ASP-792; ASP-793; LYS-1333 AND 1454-ASP-GLU-1455.
RX MEDLINE=21362977; PubMed=11469806; DOI=10.1006/abbi.2001.2441;
RA Cui L., Hou Y.-X., Riordan J.R., Chang X.-B.;
RT "Mutations of the Walker B motif in the first nucleotide binding
domain of multidrug resistance protein MRP1 prevent conformational
maturation."
RL Arch. Biochem. Biophys. 392:153-161 (2001).
RN [10]
RP MUTAGENESIS OF TRP-1246.
RX MEDLINE=21238219; PubMed=11278867; DOI=10.1074/jbc.M011246200;
RA Ito K., Olsen S.L., Qiu W., Deeley R.G., Cole S.P.C.;
RT "Mutation of a single conserved tryptophan in multidrug resistance
protein 1 (MRP1/ABCC1) results in loss of drug resistance and
selective loss of organic anion transport."
RL J. Biol. Chem. 276:15616-15624 (2001).
RN [11]
RP VARIANTS GLN-633 AND VAL-671.
RX MEDLINE=2029630; PubMed=10835642; DOI=10.1038/76102;
RA Le Saux O., Urban Z., Tschuch C., Csizsar K., Bacchelli B.,
Quaglini D., Pasquali-Ronchetti I., Pope F.M., Richards A., Terry S.,
Bercovich L., de Paeppe A., Boyd C.D.;
RT "Mutations in a gene encoding an ABC transporter cause pseudoxanthoma
elasticum."
RL Nat. Genet. 25:223-227 (2000).
RN [12]
RP VARIANT VAL-671.
RX MEDLINE=20283940; PubMed=10811882; DOI=10.1073/pnas.100041297;
RA Ringpfeil F., Lebowitz M.G., Cristiano A.M., Uitto J.;
RT "Pseudoxanthoma elasticum: mutations in the MRP6 gene encoding a
transmembrane ATP-binding cassette (ABC) transporter."
RL Proc. Natl. Acad. Sci. U.S.A. 97:6001-6006 (2000).
RN [13]
RP VARIANT SER-433, AND CHARACTERIZATION OF VARIANT VAL-671.
RX MEDLINE=215749494; PubMed=11721885; DOI=10.1007/s100380170017;
RA Conrad S., Kaufmann H.-W., Ito K., Deeley R.G., Cole S.P.C.,
Schrenk D.;
RT "Identification of human multidrug resistance protein 1 (MRP1)
mutations and characterization of a G671V substitution."
RL J. Hum. Genet. 46:656-663 (2001).
RN [14]
RP VARIANTS MET-117 AND LEU-1512.
RX MEDLINE=20579883; PubMed=11139250;
DOI=10.1002/1098-1004(2001)17:1<74::AID-HUMU14>3.0.CO;2-F;
RA Perdut J., Germain D.P.;
RT "Identification of novel polymorphisms in the pms and MRP1 (ABCC1)
genes at locus 16p13.1 and exclusion of both genes as responsible for
pseudoxanthoma elasticum."
RL Hum. Mutat. 17:74-75 (2001).

[15]
RN VARIANTS SER-433, ILE-73; GLN-723 AND GLN-1058.
RP MEDLINE=21163848; PubMed=1126082;
RX DOI=10.1097/00008571-200103000-00008;
RA Ito S., Ito T., Tanabe M., Suzuki A., Higuchi S., Otsubo K.;
RT "Polymorphism of the ABC transporter genes, MRP1, MRP2 and MRP2/cMOAT,
in healthy Japanese subjects."
RL Pharmacogenetics 11:175-184 (2001).
RN -1- FUNCTION: May participate directly in the active transport of
drugs into subcellular organelles or influence drug distribution
indirectly.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=8;
CC Comment-Additional isoforms seem to exist. Experimental
CC confirmation may be lacking for some isoforms;
CC Name=All isoforms;
CC IsoId=P33527-1; Sequence=Displayed;
CC Name=Delexon-17;
CC IsoId=P33527-2; Sequence=VSP_000037;
CC Name=Delexon-18;
CC IsoId=P33527-3; Sequence=VSP_000038;
CC Name=Delexon-30;
CC IsoId=P33527-4; Sequence=VSP_000039;
CC Name=Delexon-17-18;
CC IsoId=P33527-5; Sequence=VSP_000037, VSP_000038;
CC Name=Delexon-17-30;
CC IsoId=P33527-6; Sequence=VSP_000037, VSP_000039;
CC Name=Delexon-18-30;
CC IsoId=P33527-7; Sequence=VSP_000038, VSP_000039;
CC Name=Delexon-17-18-30;
CC IsoId=P33527-8; Sequence=VSP_000037, VSP_000038, VSP_000039;
CC -1- TISSUE SPECIFICITY: Lung, testis and peripheral blood mononuclear
cells.
CC -1- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.
CC -1- SIMILARITY: Contains 2 ABC transmembrane type-1 domains.
CC -1- DATABASE: NME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.inbiolegen.fr/services/chromocancer/genes/MRP106.html".
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: L05628; AAB4616.1; -; mRNA.
CC EMBL: AF022853; AAB83983.1; -; Genomic DNA.
CC EMBL: AF022824; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022825; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022826; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022827; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022828; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022829; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022830; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022831; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022832; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022833; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022834; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022835; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022836; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022837; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022838; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022839; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022840; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022841; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022842; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022843; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022844; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022845; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022846; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022847; AAB83983.1; JOINED; Genomic DNA.
CC EMBL: AF022848; AAB83983.1; JOINED; Genomic DNA.

DR EMBL AF022848; AAB83979.1; JOINED; Genomic DNA.
DR EMBL AF022847; AAB83979.1; JOINED; Genomic DNA.
DR EMBL AF022846; AAB83979.1; JOINED; Genomic DNA.
DR EMBL AF022845; AAB83979.1; JOINED; Genomic DNA.
DR EMBL AF022844; AAB83979.1; JOINED; Genomic DNA.
DR EMBL AF022843; AAB83979.1; JOINED; Genomic DNA.
DR EMBL AF022842; AAB83979.1; JOINED; Genomic DNA.
DR EMBL AF022841; AAB83979.1; JOINED; Genomic DNA.
DR EMBL AF022840; AAB83979.1; JOINED; Genomic DNA.
DR EMBL AF022838; AAB83979.1; JOINED; Genomic DNA.
DR EMBL AF022836; AAB83979.1; JOINED; Genomic DNA.
DR EMBL AF022834; AAB83979.1; JOINED; Genomic DNA.
DR EMBL AF022832; AAB83979.1; JOINED; Genomic DNA.
DR EMBL AF022826; AAB83979.1; JOINED; Genomic DNA.
DR EMBL AF022825; AAB83979.1; JOINED; Genomic DNA.
DR EMBL AF022824; AAB83979.1; JOINED; Genomic DNA.
DR EMBL AF022830; AAB83979.1; JOINED; Genomic DNA.
DR HSSP; P08716; IMT0.
DR Ensembl; ENSG00000103222; Homo sapiens.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity; coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran; 2.
DR Pfam; PF00006; ABC_transporter; 2.
DR ProDom; PD000006; Abc_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS00929; ABC_TMIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR ATP-binding; Nucleotide-binding.
DR NON_TER 1
SQ SEQUENCE 1515 AA; 169853 MW; DC85592817C439FE CRC64;
Query Match 78.5%; Score 7769; DB 2; Length 1515;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 377 QTVLHGYFHI CFSVGNRIKTAIVIGAYRKALVITNSAKSSTVGEIVNIMSDAQRPM 436
DB 361 QTVLHGYFHICFVSGNRIKTAIVIGAYRKALVITNSAKSSTVGEIVNIMSDAQRPM 420
QY 437 LATIINNIWGNPLOVITLALVILMTNIGPSYLAGVAVNWVPAVNAWAMKTKTVQVAMK 496
DB 421 LATIINNIWGNPLOVITLALVILMTNIGPSYLAGVAVNWVPAVNAWAMKTKTVQVAMK 480
QY 497 SKDNRIKLMNEILNGIKVLKLYAMELAFKDVLAIRQBELKVLKKSAYLSAVGTFTWCT 556
DB 481 SKDNRIKLMNEILNGIKVLKLYAMELAFKDVLAIRQBELKVLKKSAYLSAVGTFTWCT 540
QY 557 PFLVALCTFAVYVITDENNIIDAQTAFAVSLAFNIIPLPNIIPMVISIYQASVSLKRL 616
DB 541 PFLVALCTFAVYVITDENNIIDAQTAFAVSLAFNIIPLPNIIPMVISIYQASVSLKRL 600
QY 617 RIFLSHELEPDSIERPVDGCGTNSITVRNAFTMARSPPTLNGITTSIPGALVAV 676
DB 601 RIFLSHELEPDSIERPVDGCGTNSITVRNAFTMARSPPTLNGITTSIPGALVAV 660
QY 677 VGQVCGCKSSLALSLAEMDKVEGHVAIKGSVAVYPOQAWIQNDSLRENILFGQLEBPY 736
DB 661 VGQVCGCKSSLALSLAEMDKVEGHVAIKGSVAVYPOQAWIQNDSLRENILFGQLEBPY 720
QY 737 YRSVIQACALLPDLIELPSGDRTEIGEGVNLSCGQKQKVSABAVYSNADIYLPDDPLS 796
DB 721 YRSVIQACALLPDLIELPSGDRTEIGEGVNLSCGQKQKVSABAVYSNADIYLPDDPLS 780
QY 797 AYDAHVGHRIPEENYIGPKGMKNKTRILLYTHSMGYLPQVDYIYMSGKISEMSYQELL 856
DB 781 AYDAHVGHRIPEENYIGPKGMKNKTRILLYTHSMGYLPQVDYIYMSGKISEMSYQELL 840
QY 857 ARDGAFAFLRTYASTEOBDAEENGVTGVSQPKGEAKOMENGLVYTDASAKOLQRLSS 916
DB 841 ARDGAFAFLRTYASTEOBDAEENGVTGVSQPKGEAKOMENGLVYTDASAKOLQRLSS 900
QY 917 SSSYSGDISSRHNSTABLOKAEAKKETWKLMEADKAQTGVKLSYWDYKAIQGLFISF 976
DB 901 SSSYSGDISSRHNSTABLOKAEAKKETWKLMEADKAQTGVKLSYWDYKAIQGLFISF 960
QY 977 LSTIFLPMCNHVSALASNYWLSLWTDPIVNGTQHTKRLSYGALGISQGIAPVGYMA 1036
DB 961 LSTIFLPMCNHVSALASNYWLSLWTDPIVNGTQHTKRLSYGALGISQGIAPVGYMA 1020
QY 1037 VSIGIILASRCLAHVDLHSILRSPMSFPERPTSGNLVNRFSKELDTYVDSMIPVYKPMFG 1096
DB 1021 VSIGIILASRCLAHVDLHSILRSPMSFPERPTSGNLVNRFSKELDTYVDSMIPVYKPMFG 1080
QY 1097 SLFNVIGACIVILLATPIAIIIPPLGLIYFVQRFYVASSRQLKRLSVSRSPYSHFN 1156
DB 1081 SLFNVIGACIVILLATPIAIIIPPLGLIYFVQRFYVASSRQLKRLSVSRSPYSHFN 1140
QY 1157 ETLIGVSTIRAFEOERFIHOSDLKVDENQAYYPSIYANWLAVRLECNGNCIYLPAL 1216
DB 1141 ETLIGVSTIRAFEOERFIHOSDLKVDENQAYYPSIYANWLAVRLECNGNCIYLPAL 1200
QY 1217 FAVISRHSLSAGLVGLSVSYLQVTTYNTYNTVVRMSSEMETNIIVAVERLKEYSSETEKAPW 1276
DB 1201 FAVISRHSLSAGLVGLSVSYLQVTTYNTYNTVVRMSSEMETNIIVAVERLKEYSSETEKAPW 1260
QY 1277 QIOETAPPSWPOVGRVFRNYCLRYREDLDFVLRIHIVITINGEKVGI VGRTAGKSSL 1336
DB 1261 QIOETAPPSWPOVGRVFRNYCLRYREDLDFVLRIHIVITINGEKVGI VGRTAGKSSL 1320
QY 1337 TLGLFRINESABGIIIDGINIAKIGLHDLRFKTTIIIPQDVLRSGLIRNMLDPPSQYSD 1396
DB 1321 TLGLFRINESABGIIIDGINIAKIGLHDLRFKTTIIIPQDVLRSGLIRNMLDPPSQYSD 1380
QY 1397 EEVWTSLELAHKKQVPSALPKDLHECABGGENISVQGRQVLCARALLRKTKIIVLDEA 1456
DB 1381 EEVWTSLELAHKKQVPSALPKDLHECABGGENISVQGRQVLCARALLRKTKIIVLDEA 1440
QY 1457 TAAVDLLETDDLIOSTIRTOFEDCTVLTIARHLNLTIMDYTRVIVLDKGEIOEYGA PSDLLQ 1516

Db 1441 TAAVLDFTDDLIQSTIFRQFEDCTVLTIAHLRLNTMDTRVIVLDKGEIQEIGAPSDLLQ 1500
 QY 1517 ORGLFYSAKXAGLV 1531
 Db 1501 QRGLFYSAKXAGLV 1515

RESULT 3

0864R9_MACPA PRELIMINARY; PRT; 1531 AA.
 ID 0864R9_MACPA
 AC 0864R9;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Multidrug resistance protein 1B.
 OS Name=Mppl.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 OC Cercopithecidae; Cercopithecinae; Macaca.
 OC NCBI_TaxID=9541;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22544876; PubMed=12657726;
 RA Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,
 RT "Cloning and functional characterization of the multidrug resistance-
 associated protein (MRP1/ABCC1) from the cynomolgus monkey";
 RT Mol. Cancer Ther. 2:307-316(2003).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.

RA Perry W.L. III, Godinot N.; EMBL/GenBank/DBJ databases.
 RL Submitted (Aug-2002) to the ABC transporter family.
 CC -1- SIMILARITY: Belongs to the ABC transporter family.
 DR EMBL; AY146673; AAN65349.1; -, mRNA.
 DR HSSP; P08716; IMTO.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005524; P:ATP binding; IEA.
 DR GO; GO:0016887; F:ATPase activity; IEA.
 DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
 DR GO; GO:0000166; F:nucleotide binding; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR011527; ABC_membrane_1.
 DR InterPro; IPR001140; ABC_TM_transp.
 DR InterPro; IPR003439; ABC_transp_like.
 DR InterPro; IPR005292; MRP_assoc.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00664; ABC_membrane; 2.
 DR Pfam; PF00005; ABC_tran; 2.
 DR ProDom; PD000006; ABC_transporter; 2.
 DR SMART; SM00382; AAA; 2.
 DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
 DR PROSITE; PS00929; ABC_TMIF; 2.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
 DR PROSITE; PS08893; ABC_TRANSPORTER_2; 2.
 DR PROSITE; PS0107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR AP-binding; Nucleotide-binding; Repeat.
 KW SEQUENCE 1531 AA; 171659 MW; 1AE788BFD9E9459 CRC64;

Query Match 77.7%; Score 7695; DB 2; Length 1531;
 Best Local Similarity 97.5%; Pred. No. 0;
 Matches 1493; Conservative 20; Mismatches 18; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTVNTSNPDTKCFQNTVTVVWPCFYLMACFPFYLYSRH 60
 Db 1 MALRGFCSADGSDPLMDMNTVNTSNPDTKCFQNTVTVVWPCFYLMACFPFYLYSRH 60
 QY 61 DRGVIQMTPLNKTALIGFLIMIVCANADLFYSFWEERSRGIFLAVPLVSPITLIGITTLA 120
 Db 61 DRGVIQMTPLNKTALIGFLIMIVCANADLFYSFWEERSRGIFLAVPLVSPITLIGITTLA 120

QY 121 TFLIQLERRKVOSSGIMLTFLMALVCAALALIRSKIMTALKEADQVDFRDIIFYVES 180
 Db 121 TFLIQLERRKVOSSGIMLTFLMALVCAALALIRSKIMTALKEADQVDFRDIIFYVES 180
 QY 181 LLLIQVLVSCPSDRSPLFSETHIDPNCPRESSASFLSRITFWMTGLIVRGYROPLESGD 240
 Db 181 LLLIQVLVSCPSDRSPLFSETHIDPNCPRESSASFLSRITFWMTGLIVRGYROPLESGD 240
 QY 241 LMSLNKEDTSEQVIVLVKMKKCAKTRKOPVRYVYSSKDPAPKSSSKVDANEVEAL 300
 Db 241 LMSLNKEDTSEQVIVLVKMKKCAKTRKOPVRYVYSSKDPAPKSSSKVDANEVEAL 300
 QY 301 IVKSPQKEMNPSLKFVLYKTGPFLMSFFFKALHDLMMFSGPEILKLLINFDVDTKAPD 360
 Db 301 IVKSPQKEMNPSLKFVLYKTGPFLMSFFFKALHDLMMFSGPEILKLLINFDVDTKAPD 360
 QY 361 WQGYFYTLVFLVTAQTLVLYHQFPHICFVSGMKIKTAVIGAVRKALVITNSARKSSTV 420
 Db 361 WQGYFYTLVFLVTAQTLVLYHQFPHICFVSGMKIKTAVIGAVRKALVITNSARKSSTV 420
 QY 421 GEIVNLMASVDAQRPMDLATTYINMTWSAPLOYTALYILMLNGPSVLAGAVMYLVMPVN 480
 Db 421 GEIVNLMASVDAQRPMDLATTYINMTWSAPLOYTALYILMLNGPSVLAGAVMYLVMPVN 480
 QY 481 AVMAKTKTQVAMKSKDNRIKIMNELINGIKLYKLYAMELAKDKVLAIRQELKYLK 540
 Db 481 AVMAKTKTQVAMKSKDNRIKIMNELINGIKLYKLYAMELAKDKVLAIRQELKYLK 540
 QY 541 KSAVLASAVGTFTWCTPFVVALCTFAVYVITIDENNILDAQAFSLAFNLRPPLNLP 600
 Db 541 KSAVLASAVGTFTWCTPFVVALCTFAVYVITIDENNILDAQAFSLAFNLRPPLNLP 600
 QY 601 MVISSIVQASVSLRLRILFISHEELBPSDIERRPVKDGGTNSITVENATFTMARSDPT 660
 Db 601 MVISSIVQASVSLRLRILFISHEELBPSDIERRPVKDGGTNSITVENATFTMARSDPT 660
 QY 661 LINGITFIPRGALVAVVGVCGGKSSLSALLAEMDVEGHVALKGSVAVYPOQAMQND 720
 Db 661 LINGITFIPRGALVAVVGVCGGKSSLSALLAEMDVEGHVALKGSVAVYPOQAMQND 720
 QY 721 SLRENIILFGQLEBPYRSYIQAALPDLIELPSGRTIEGKGVNLGGQKQKRVSLAR 780
 Db 721 SLRENIILFGQLEBPYRSYIQAALPDLIELPSGRTIEGKGVNLGGQKQKRVSLAR 780
 QY 781 AVYSNADIVYFDPLSAVDAAVGHGHIPEENVIGPKMKLKNKTRILLVTHSMGYLPQVDYIIV 840
 Db 781 AVYSNADIVYFDPLSAVDAAVGHGHIPEENVIGPKMKLKNKTRILLVTHSMGYLPQVDYIIV 840
 QY 841 MSGGKISEMGSYQELLARDGAFAPFLRTVYSAEQEODPENGVGVSGPKKAKQKMG 900
 Db 841 MSGGKISEMGSYQELLARDGAFAPFLRTVYSAEQEODPENGVGVSGPKKAKQKMG 900
 QY 901 LVYDSAGKQOROLSSSSSYSGDISRRHNSTAELQKAEAKKEFTKMLEADKQOTGVKL 960
 Db 901 LVYDSAGKQOROLSSSSSYSGDISRRHNSTAELQKAEAKKEFTKMLEADKQOTGVKL 960
 QY 961 SVYWDYVKAIGLIFSLIFSLIFLFCNHYAALASNYWLSLWTDPIVNGTQHTKRVLSVYG 1020
 Db 961 SVYWDYVKAIGLIFSLIFSLIFLFCNHYAALASNYWLSLWTDPIVNGTQHTKRVLSVYG 1020
 QY 1021 ALGISQGIANFGYSMAVSTIGIILASRCLAHVDLHSILRSFMSFPERPSGNLVNRSKEL 1080
 Db 1021 ALGISQGIANFGYSMAVSTIGIILASRCLAHVDLHSILRSFMSFPERPSGNLVNRSKEL 1080
 QY 1081 DTVDSMIPVYIKMFMGSLFNVIGACIVYLLATPLAIIIPPLGLIYEFVORFYVASSROL 1140
 Db 1081 DTVDSMIPVYIKMFMGSLFNVIGACIVYLLATPLAIIIPPLGLIYEFVORFYVASSROL 1140
 QY 1141 KRLESVSRSPVYSHFNETLIGSVIRAFEEQEPFIHQSDIKVDENQKAYYPSIVANRWLA 1200
 Db 1141 KRLESVSRSPVYSHFNETLIGSVIRAFEEQEPFIHQSDIKVDENQKAYYPSIVANRWLA 1200
 QY 1201 VRLCEVGNCLVFAALFAVISRSLASGLVGLSVSYSLQVTTYLLNMLVRRSMENETIYA 1260

Db 1201 VRLCEVGCIVLFAALFAVIRSHSISAGLVLSYSIQVTTYINMLVMSSEMETNIVA 1260
QY 1261 VERKEVSETEKEAPMOQIOETAPSPWQVGRVERRNCIAYRELDIVLPHINVTNGG 1320
Db 1261 VERKEVSETEKEAPMOQIOETAPSPWQVGRVERRNCIAYRELDIVLPHINVTNGG 1320
QY 1321 EKVGI VGTGAGKSLTGLFRINSEAGEIIIDGINIAKIGLHDLRKFITIIIPQDPVLF 1380
Db 1321 EKVGI VGTGAGKSLTGLFRINSEAGEIIIDGINIAKIGLHDLRKFITIIIPQDPVLF 1380
QY 1381 SGSI RPNMIDPFSQYSDEEVTWTSLEIAHLKDPYSALPDKLDEHCAEGENLSVGQRQVCL 1440
Db 1381 SGSI RPNMIDPFSQYSDEEVTWTSLEIAHLKDPYSALPDKLDEHCAEGENLSVGQRQVCL 1440
QY 1441 ARALLRKTIIVLEBATAVADLETDLLQSTIRIQFEDCTVLTIAHRLNTIMDYTRIVL 1500
Db 1441 ARALLRKTIIVLEBATAVADLETDLLQSTIRIQFEDCTVLTIAHRLNTIMDYTRIVL 1500
QY 1501 DKGEIOEYGAPSDLLQGRGLFYMAKADGLV 1531
Db 1501 DKGEIOEYGAPSDLLQGRGLFYMAKADGLV 1531

RESULT 4

Q864S0 MACFA PRELIMINARY; PRT; 1531 AA.

AC 0864S0
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Multidrug resistance protein 1A.
GN Name=MRP1;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopitheciidae; Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RX MEDLINE=22544876; PubMed=12657726;
RA Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,
RA Dantzig A.H., Perry W.L.;
RT "Cloning and functional characterization of the multidrug resistance-
RT associated protein (MRP1/ABCC1) from the cynomolgus monkey.",
RL Mol. Cancer Ther. 2:307-316(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RA Perry W.L., Ili, Godinot N.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: belongs to the ABC transporter family.
EMBL: AY146672; AAN65348.1; -; mRNA.
DR HSSP; P08716; 1MT0.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran_2.
DR ProDom; PD000006; ABC_transporter_2.
DR SMART; SM00382; AAA_2.
DR TIGFams; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TM1F; 2.

DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN KINASE_ATP; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1531 AA; 171701 MW; 168712B8CC2D2B89 CRC64;

Query Match 77.6%; Score 7689; DB 2; Length 1531;

Best Local Similarity 97.6%; Pred. No. 0;
Matches 1494; Conservative 18; Mismatches 19; Indels 0; Gaps 0;

QY 1 MALRGFSGADSDLDLMMNNTWNTNSNDFKCPQNTLVWPCRYLWACPFYFLYLSRH 60
Db 1 MALRGFSGADSDLDLMMNNTWNTNSNDFKCPQNTLVWPCRYLWACPFYFLYLSRH 60
QY DRGYIOMTPLNKTKTALGFLLMIVCMADLFSPFERSRGIFLAVFLVSPFLIGITLLA 120
Db DRGYIOMTPLNKTKTALGFLLMIVCMADLFSPFERSRGIFLAVFLVSPFLIGITLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFWLVAVCALILRSKIMTALKEDAQVDLFRDITFYVFS 180
Db 121 TFLIQLERRKGVSSGIMLTFWLVAVCALILRSKIMTALKEDAQVDLFRDITFYVFS 180
QY 121 TFLIQLERRKGVSSGIMLTFWLVAVCALILRSKIMTALKEDAQVDLFRDITFYVFS 180
Db 121 TFLIQLERRKGVSSGIMLTFWLVAVCALILRSKIMTALKEDAQVDLFRDITFYVFS 180
QY 181 LLLIQVLVSCFSDSPLEFSETIHDNPNCPRESSASFLSRTIFPMWITGLIVRGYRQPLESSD 240
Db 181 LVLIQVLVSCFSDSPLEFSETIHDNPNCPRESSASFLSRTIFPMWITGLIVRGYRQPLESSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVVYSSKDPAOKESSKYDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVVYSSKDPAOKESSKYDANEVEAL 300
QY 301 IVKSPQKEMNSLFRKVLVKTGPYFLMSFFPKAIHDLMMFSGPOLIKLILFVNDTKAPD 360
Db 301 IVKSPQKEMNSLFRKVLVKTGPYFLMSFFPKAIHDLMMFSGPEILIKLILFVNDTKAPD 360
QY 361 WQGFYTYLLEVTACLOTVLHQQYFHI CFVSGMIRKTAIVAAVRKALVITNAKSSSTV 420
Db 361 WQGFYTYLLEVTACLOTVLHQQYFHI CFVSGMIRKTAIVAAVRKALVITNAKSSSTV 420
QY 421 GEIVNLSVDAQRFPMDLATYINMISAPLQVTLALYLMLMGPSVLAVGAVMYLMPVN 480
Db 421 GEIVNLSVDAQRFPMDLATYINMISAPLQVTLALYLMLMGPSVLAVGAVMYLMPVN 480
QY 481 AVMAKTKTYOVANMKSKDNRIKLMNBILNGIKVILKYAMELAFKDYALAIROBELKYLK 540
Db 481 AVMAKTKTYOVANMKSKDNRIKLMNBILNGIKVILKYAMELAFKDYALAIROBELKYLK 540
QY 541 KSAVLSANGFTTWCCTPFLVALCTPAVYVYTDENNILDAQAFAPSLALFNILRPPLNLP 600
Db 541 KSAVLSANGFTTWCCTPFLVALCTPAVYVYTDENNILDAQAFAPSLALFNILRPPLNLP 600
QY 601 MYISIVQASVSLKRLRIFLSHEELEPDSIRRPVKDGGTNSITVRNATFTWASDPPT 660
Db 601 MYISIVQASVSLKRLRIFLSHEELEPDSIRRPVKDGGTNSITVRNATFTWASDPPT 660
QY 661 LINGITFSIPBGALVAVVQVCGKSSILSALLAEMDKVEGHVAIKGSVAVYVPOQAMIQND 720
Db 661 LINGITFSIPBGALVAVVQVCGKSSILSALLAEMDKVEGHVAIKGSVAVYVPOQAMIQND 720
QY 721 SLRENTILFGCQLEBYRHSVYQAALLPDLLELTPSGDRTIEGKGVNLSGGQKQKRVSLAR 780
Db 721 SLOENTILFGCQLEBYRHSVYQAALLPDLLELTPSGDRTIEGKGVNLSGGQKQKRVSLAR 780
QY 781 AVYSNADLYLPDDPLSAVDAVGHKIFENVIGPKMLNKTRILVTHSMSYLPOVDVILV 840
Db 781 AVYCNADLYLPDDPLSAVDAVGHKIFENVIGPKMLNKTRILVTHSMSYLPOVDVILV 840
QY 841 MSGKISMSGVQYOBELARDGAFABFLRTYASTEDQDAENGVTGVSFGKAKEMKMG 900
Db 841 MSGKISMSGVQYOBELARDGAFABFLRTYASTEDQDAENGVTGVSFGKAKEMKMG 900
QY 901 LVTTDSAGKOLOROLSSSSSSSGDLSRHNSITAELOKAKKBEYTKLMEADKAOQGOYKL 960
Db 901 LVTTDSAGKOLOROLSSSSSSSGDLSRHNSITAELOKAKKBEYTKLMEADKAOQGOYKL 960

```
QY 961 SVYWDYKAIIGLFTISFLISIFLEMCNHSALASNYWLSLTWDDPTVNGTQEHKRLSYVG 1020
Db 961 SVYWDYKAIIGLFTISFLISIFLEMCNHSALASNYWLSLTWDDPTVNGTQEHKRLSYVG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGGILASRLYHVDLHLSILSPSPSPFPERRPSGNLVNRFSEKL 1080
Db 1021 ALGISOGIAVFGYSMAVSIIGGILASRLYHVDLHLSILSPSPSPFPERRPSGNLVNRFSEKL 1080
QY 1081 DTYSMTPEVTKMFMGSLFNVIACIVILATPTIAIIPPLGLIYFVQRFYVASSQOL 1140
Db 1081 DTYSMTPEVTKMFMGSLFNVIACIVILATPTIAIIPPLGLIYFVQRFYVASSQOL 1140
QY 1141 KRLSVSRSPVYSHPHNETLGVSVIRAFEEQERFIHOSDLKYDENOKAYYSIVANRMLA 1200
Db 1141 KRLSVSRSPVYSHPHNETLGVSVIRAFEEQERFIHOSDLKYDENOKAYYSIVANRMLA 1200
QY 1201 VRLFCVGNICIVLPAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSSEMETNIVA 1260
Db 1201 VRLFCVGNICIVLPAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSSEMETNIVA 1260
QY 1261 VERLKESETTEKAPWQIOETAPPSNPQVGRVEFRNYCLRYREDLDFVLRHINVTINGG 1320
Db 1261 VERLKESETTEKAPWQIOETAPPSNPQVGRVEFRNYCLRYREDLDFVLRHINVTINGG 1320
QY 1321 EKVGIVERTGAGKSLTLGLFRINESAGEIITIDGINAKTGLADLPKRTITIPQDPYLF 1380
Db 1321 EKVGIVERTGAGKSLTLGLFRINESAGEIITIDGINAKTGLADLPKRTITIPQDPYLF 1380
QY 1381 SGSLRNMLDPEFSQYSDSEVWTSLELAHLKDFVSALPKDLHBCAEGENLSVQORQVCL 1440
Db 1381 SGSLRNMLDPEFSQYSDSEVWTSLELAHLKDFVSALPKDLHBCAEGENLSVQORQVCL 1440
QY 1441 ARALLRTKTLIVDEATPAVDLETFDILQSTIRFQFEDCTVLTAAHRLNTIMDYTRYIVL 1500
Db 1441 ARALLRTKTLIVDEATPAVDLETFDILQSTIRFQFEDCTVLTAAHRLNTIMDYTRYIVL 1500
QY 1501 DKGEIOEYGA PSDLLQORGLFYMAKXAGLV 1531
Db 1501 DKGEIOEYGA PSDLLQORGLFYMAKXAGLV 1531

RESULT 5
Q9U097_HUMAN PRELIMINARY; PRT; 1459 AA.
ID Q9U097_HUMAN PRELIMINARY; PRT; 1459 AA.
AC Q9U097;
DT 01-MAY-2000 (Tremblrel, 13, Created)
DT 01-MAY-2000 (Tremblrel, 13, Last sequence update)
DT 01-MAR-2004 (Tremblrel, 26, Last annotation update)
DE Multidrug resistance protein (Fragment).
GN Name=MRP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;
OC Homo.
NCBI_TaxID=9606;
RX NCBIOTIDE SEQUENCE.
RA Grant C.B., Kurz E.U., Cole S.P.C., Deeley R.G.;
RT "Analysis of the intron-exon organization of the human multidrug-
RT resistance protein gene (MRP) and alternative splicing of its mRNA.";
RL Genomic 45368-37611977).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF022838; AAB83980.1; -; Genomic DNA.
DR EMBL; AF022834; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022825; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022826; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022828; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022830; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022832; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022834; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022836; AAB83980.1; JOINED; Genomic DNA.
```

```
DR EMBL; AF022838; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022848; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022847; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022846; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022845; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022844; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022843; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022842; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022841; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022839; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022835; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022851; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022850; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022849; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022837; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022835; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022833; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022831; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022829; AAB83980.1; JOINED; Genomic DNA.
DR EMBL; AF022827; AAB83980.1; JOINED; Genomic DNA.
DR HSDP; P08716; 1MT0.
DR Ensemble; ENSG00000103222; Homo sapiens.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR011527; ABC membrane 1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_TM_transp.
DR InterPro; IPR005292; MRP_assoc.
DR InterPro; IPR007119; Prot_kinase.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA_2.
DR TIGRfams; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PSS0929; ABC_TM1F; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN_1.
DR PROSITE; PSS0893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR ATP-binding; Nucleotide-binding.
DR NON_TER 1
SQ SEQUENCE 1459 AA; 16381 MW; 6A016A5A6E7A95C CRC64;

Query Match 75.2%; Score 7444; DB 2; Length 1459;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1459; Conservative 0; Mismatches 0; Indels 56; Gaps 1;

QY 17 DNVVTWNTSNPDEFTKCFQNTLVVWVPCFYLMACFPFYFLYLSRHDRGYIOMTPIANKTKTA 76
Db 1 DNVVTWNTSNPDEFTKCFQNTLVVWVPCFYLMACFPFYFLYLSRHDRGYIOMTPIANKTKTA 76
QY 77 LGFLMIVCAADLFYSWERSRGIFLAPVFLVSPPTLIGITTLATFLIQERRKGVQSSG 136
Db 77 LGFLMIVCAADLFYSWERSRGIFLAPVFLVSPPTLIGITTLATFLIQERRKGVQSSG 136
QY 61 LGFLMIVCAADLFYSWERSRGIFLAPVFLVSPPTLIGITTLATFLIQERRKGVQSSG 120
Db 61 LGFLMIVCAADLFYSWERSRGIFLAPVFLVSPPTLIGITTLATFLIQERRKGVQSSG 120
QY 137 IMLTFLMVALVCAALILRSKIMTALKEBDQVDLFRDITFVYVSSLILQVLSCFSPDRSP 196
Db 137 IMLTFLMVALVCAALILRSKIMTALKEBDQVDLFRDITFVYVSSLILQVLSCFSPDRSP 196
QY 121 IMLTFLMVALVCAALILRSKIMTALKEBDQVDLFRDITFVYVSSLILQVLSCFSPDRSP 180
Db 121 IMLTFLMVALVCAALILRSKIMTALKEBDQVDLFRDITFVYVSSLILQVLSCFSPDRSP 180
QY 197 LFSETHDPNPCESSASFSLRITFWITGLIVRGYROPLEGSDLSLNKEDTSEQVVPV 256
Db 197 LFSETHDPNPCESSASFSLRITFWITGLIVRGYROPLEGSDLSLNKEDTSEQVVPV 256
QY 181 LFSETHDPNPCESSASFSLRITFWITGLIVRGYROPLEGSDLSLNKEDTSEQVVPV 240
Db 181 LFSETHDPNPCESSASFSLRITFWITGLIVRGYROPLEGSDLSLNKEDTSEQVVPV 240
QY 257 LVNMWKECKTKTKQPVYVYSSKDPAPQKESKVDANEVEALLYKSPKKNPSPSLFKV 316
Db 257 LVNMWKECKTKTKQPVYVYSSKDPAPQKESKVDANEVEALLYKSPKKNPSPSLFKV 316
QY 241 LVNMWKECKTKTKQPVYVYSSKDPAPQKESKVDANEVEALLYKSPKKNPSPSLFKV 300
Db 241 LVNMWKECKTKTKQPVYVYSSKDPAPQKESKVDANEVEALLYKSPKKNPSPSLFKV 300
QY 317 LVYTPGPYFAMSFPAIHDLMMFSGPOLIKLIKFNTPDKADMGQVFTYVTLPTATCL 376
Db 317 LVYTPGPYFAMSFPAIHDLMMFSGPOLIKLIKFNTPDKADMGQVFTYVTLPTATCL 376
```

Db 301 LYKTEGPEYELMSPEFFKAIDHLMFSGPOLIKLLIKFVNDTAPDMOGYFYTVLLFVTACL 360
Qy 377 QTVLHGOYFHIQFVSGMRIKTAVIGAVYRKALVITNSARKSSTVEIYNLMSVDAQRPMD 436
Db 361 QTVLHGOYFHIQFVSGMRIKTAVIGAVYRKALVITNSARKSSTVEIYNLMSVDAQRPMD 420
Qy 437 LATYINMWSAPLOYLIALYLLMLNGPSVLAGVAVMYLWVPVNAVMAMKTKTYQVAHMK 496
Db 421 LATYINMWSAPLOYLIALYLLMLNGPSVLAGVAVMYLWVPVNAVMAMKTKTYQVAHMK 480
Qy 497 SKDMRKIKMNEELINGIKYLKIYAMELAFKDYALROBELKYLKKSATLSAVGFTWCT 556
Db 481 SKDMRKIKMNEELINGIKYLKIYAMELAFKDYALROBELKYLKKSATLSAVGFTWCT 540
Qy 557 PFLVALCTFAYVYVITDENNIIDAQTAFAVSLAFNLRPLNLLPWVSIIVQASVSLRL 616
Db 541 PFLVALCTFAYVYVITDENNIIDAQTAFAVSLAFNLRPLNLLPWVSIIVQASVSLRL 600
Qy 617 RIFLSHEELEPDSIRRRPVKGGGTNSITVRNATFTWASDDPTLNGITFSPBGALVAV 676
Db 601 RIFLSHEELEPDSIRRRPVKGGGTNSITVRNATFTWASDDPTLNGITFSPBGALVAV 660
Qy 677 VGVQCCGSSLSLALLAEMDKVEGHVATKGSVAVYPOQAMTQNDSLRENILLPGCOLERP 736
Db 661 VGVQCCGSSLSLALLAEMDKVEGHVATKGSVAVYPOQAMTQNDSLRENILLPGCOLERP 720
Qy 737 YRSVIOACALLPDEILPBGDRTEIGEGKVNLSGGQXQVSLARAVYNSADIYLFDDPLS 796
Db 721 YRSVIOACALLPDEILPBGDRTEIGEGKVNLSGGQXQVSLARAVYNSADIYLFDDPLS 747
Qy 797 AVDAHVGRIFFENVIGPKMLKNKTRIIIVTHSMSTLPQVDVITVMSGGKISMSGYOELL 856
Db 748 -----KTRIIIVTHSMSTLPQVDVITVMSGGKISMSGYOELL 784
Qy 857 ARDGAFAEFLRTYATSTEOBODAEENGVTGVSQPGKEAKOMENGMVITQSAGQOLROJSS 916
Db 785 ARDGAFAEFLRTYATSTEOBODAEENGVTGVSQPGKEAKOMENGMVITQSAGQOLROJSS 844
Qy 917 SSSYSGDLSRRHNSYAELOKAEKKEETWKLMEARKAQTQYKLSVYMDYMKALGLFISF 976
Db 845 SSSYSGDLSRRHNSYAELOKAEKKEETWKLMEARKAQTQYKLSVYMDYMKALGLFISF 904
Qy 977 LSLIFLMCNHVSALASNTWLSLMTDDPIVNGTQEHKVLASYGALGISQGIATVFGYSMA 1036
Db 905 LSLIFLMCNHVSALASNTWLSLMTDDPIVNGTQEHKVLASYGALGISQGIATVFGYSMA 964
Qy 1037 VSIQGIILASRCLAHVLLHSILASPMSPFERTSGMLVNRFSKELDTVDSMIPEVIMKFMG 1096
Db 965 VSIQGIILASRCLAHVLLHSILASPMSPFERTSGMLVNRFSKELDTVDSMIPEVIMKFMG 1024
Qy 1097 SLFENVIGACIVLLATPIAAITIIIPPLGLIYFPQRFYVYVASSHQRLKRLSVSRSPYSHFN 1156
Db 1025 SLFENVIGACIVLLATPIAAITIIIPPLGLIYFPQRFYVYVASSHQRLKRLSVSRSPYSHFN 1084
Qy 1157 ETLGLVSVIRAEEOERFIHQSDLVKVENOKAYPSIVANRWLAVALBECVAGCIVLFAAL 1216
Db 1085 ETLGLVSVIRAEEOERFIHQSDLVKVENOKAYPSIVANRWLAVALBECVAGCIVLFAAL 1144
Qy 1217 FAVISRHSISAGLVGLSVYSLSQVTTYLNMVLRMSSEMETNIYAVBRLEKYSETEKEAPW 1276
Db 1145 FAVISRHSISAGLVGLSVYSLSQVTTYLNMVLRMSSEMETNIYAVBRLEKYSETEKEAPW 1204
Qy 1277 QIQETAPPSWPFQVQVREFRNYCLRRYREDLDVLRHINVTINGEGEVGIVGTGAGKSSL 1336
Db 1205 QIQETAPPSWPFQVQVREFRNYCLRRYREDLDVLRHINVTINGEGEVGIVGTGAGKSSL 1264
Qy 1337 TLGLRFINESAGEIIIDGINIAKIGLHDLRKITIIIPQDPVLFSGSLRMLNDPFSQYSD 1396
Db 1265 TLGLRFINESAGEIIIDGINIAKIGLHDLRKITIIIPQDPVLFSGSLRMLNDPFSQYSD 1324
Qy 1397 BEVWMTSLBLAHKDYVSALPDKLDHCEAGGENTLSVGQOLVCLARALLRKTKIIVLDBA 1456
Db 1325 BEVWMTSLBLAHKDYVSALPDKLDHCEAGGENTLSVGQOLVCLARALLRKTKIIVLDBA 1384

Qy 1457 TAAVDLETDLLIGSTIRTOEEDCTVLTIAHRLNTIMDYTRVILDKGEIOEYGAPSDILQ 1516
Db 1385 TAAVDLETDLLIGSTIRTOEEDCTVLTIAHRLNTIMDYTRVILDKGEIOEYGAPSDILQ 1444
Qy 1517 ORGLFYGMADQAGLV 1531
Db 1445 ORGLFYGMADQAGLV 1459
RESULT 6
Q9UQA0 HUMAN
ID Q9UQA0 HUMAN PRELIMINARY; PRT; 1456 AA.
AC Q9UQA0;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Multidrug resistance protein (Fragment).
GN Name=MRP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
CC NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98008927; PubMed=344662; DOI=10.1006/gene.1997.4950;
RA Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;
RT "Analysis of the intron-exon organization of the human multidrug-
resistance protein gene (MRP) and alternative splicing of its mRNA.";
RL Genomics 45:368-378(1997).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF022853; AAB83981.1; -; Genomic DNA.
DR EMBL; AF022853; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022825; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022825; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022827; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022829; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022831; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022833; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022835; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022837; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022847; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022846; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022845; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022844; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022843; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022842; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022841; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022840; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022838; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022851; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022850; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022849; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022848; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022836; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022834; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022832; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022830; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022828; AAB83981.1; JOINED; Genomic DNA.
DR HSSP; P08716; IMTO.
DR Ensembl; ENSG00000103222; Homo sapiens.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transpc.
DR InterPro; IPR003439; ABC_tranep_like.

```

DR InterPro: IPR005292; MRP assoc.
DR InterPro: IPR000719; Prot. kinase.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD00006; ABC_transporter; 2.
DR SMART: SM00382; AAA; 2.
DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
DR PROSITE: PS00929; ABC_TMIF; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE: PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR ATP-binding; Nucleotide-binding.
FT NON TER
SQ SEQUENCE 1456 AA; 163233 MW; 8D88AB22BC481FC CRC64;

Query Match 74.9%; Score 7419.5; DB 2; Length 1456;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 1456; Conservative 0; Mismatches 0; Indels 59; Gaps 1;

QY 17 DMNVTMTSNPDKFCQNTVTVVWPCFYLMACFPFYLYSRDRGTYQMTPLNKTKTA 76
DB 1 DMNVTMTSNPDKFCQNTVTVVWPCFYLMACFPFYLYSRDRGTYQMTPLNKTKTA 60

QY 77 LGFLMTVCMADLFYSFWEISRGIFLAFLVSPFLIGITTLATFLIQLERRKGVSSG 136
DB 61 LGFLMTVCMADLFYSFWEISRGIFLAFLVSPFLIGITTLATFLIQLERRKGVSSG 120

QY 137 IMLTFWLVAVCALIIRSKIMTAKEDAQVDLFRDITFYVYFSLIIQLVLSGFSRSP 196
DB 121 IMLTFWLVAVCALIIRSKIMTAKEDAQVDLFRDITFYVYFSLIIQLVLSGFSRSP 180

QY 197 LFSITHDNPCESSASFLSRITFWMTGLIVRGYRQPLEGSLMSLNKEDTSEQVVPV 256
DB 181 LFSITHDNPCESSASFLSRITFWMTGLIVRGYRQPLEGSLMSLNKEDTSEQVVPV 240

QY 257 LVKMKKECAKTRKOPKVVVSSKDPAPKESKVDANEVEALIVSPKEMPSLFKV 316
DB 241 LVKMKKECAKTRKOPKVVVSSKDPAPKESKVDANEVEALIVSPKEMPSLFKV 300

QY 317 LYKTFGEYFLMSFPFKAIDHLMFSGPQILKLIKPVNDYKAPDMOGYFTVLLFVTA 376
DB 301 LYKTFGEYFLMSFPFKAIDHLMFSGPQILKLIKPVNDYKAPDMOGYFTVLLFVTA 360

QY 377 QTVLHQYFHCIFSGRIKTAIVGAYYRKALVITNSRKSTYGEIYNLMSYDAQRMD 436
DB 361 QTVLHQYFHCIFSGRIKTAIVGAYYRKALVITNSRKSTYGEIYNLMSYDAQRMD 420

QY 437 LATYINNIWSAPLOVIALYLLMLNGPSVLAVGAVVWLVAVVAVVAMKTKTYQVAHMK 496
DB 421 LATYINNIWSAPLOVIALYLLMLNGPSVLAVGAVVWLVAVVAVVAMKTKTYQVAHMK 480

QY 497 SKDNRIKLMEILNGIKVLKLYAMELAFKDKVLAIROBELVKLKKSAVLSAVGTFTWCT 556
DB 481 SKDNRIKLMEILNGIKVLKLYAMELAFKDKVLAIROBELVKLKKSAVLSAVGTFTWCT 540

QY 557 PELVALCTPAAVYVITDENNIIDAOATAPYSLAFNILRPLNLTLMVYSSIVQASVSLKRL 616
DB 541 PELVALCTPAAVYVITDENNIIDAOATAPYSLAFNILRPLNLTLMVYSSIVQASVSLKRL 600

QY 617 RIFLSHEELEPDSIERRPVKDGGGNSITVNAFTWARSDPPTINGITSPGALVAV 676
DB 601 RIFLSHEELEPDSIERRPVKDGGGNSITVNAFTWARSDPPTINGITSPGALVAV 660

QY 677 VGVGVGCKSSILSALLAEMDKVEGHVAIKGSVAVYVPOQAWIQNDSLRENIILFGQLEBPY 736
DB 661 VGVGVGCKSSILSALLAEMDKVEGHVAIKGSVAVYVPOQAWIQNDSLRENIILFGQLEBPY 688

QY 737 YRSVYQACALLPDLIELPSGDRTEIGEGVNLSSGQKORVSLABAVYSNADIYLFDDPLS 796
DB 689 YRSVYQACALLPDLIELPSGDRTEIGEGVNLSSGQKORVSLABAVYSNADIYLFDDPLS 721

QY 797 AVDAVGHKILFENVIGPKMLKNKTRILVTHSMGYLPQVDVYIWMGSGKISEMGSYOELL 856
DB 797 AVDAVGHKILFENVIGPKMLKNKTRILVTHSMGYLPQVDVYIWMGSGKISEMGSYOELL 856

```

```

DB 722 AVDAVGHKILFENVIGPKMLKNKTRILVTHSMGYLPQVDVYIWMGSGKISEMGSYOELL 781
QY 857 ARDAFAEFLRTYASTQEQDABENGYTGYSGPCKEAKQENGMVLYDSAGKQIQROLS 916
DB 782 ARDAFAEFLRTYASTQEQDABENGYTGYSGPCKEAKQENGMVLYDSAGKQIQROLS 841
QY 917 SSSYSGISRHNNSTAELOKAEAKKETWKLMEADKQOTQGVKLSVYWDYKAIQGLFISF 976
DB 842 SSSYSGISRHNNSTAELOKAEAKKETWKLMEADKQOTQGVKLSVYWDYKAIQGLFISF 901
QY 977 LSLFLFMCNHNVSALASNYWMLSLWTDPIVNGQDHTVRVLSVYALGISQIAVFGYSMA 1036
DB 902 LSLFLFMCNHNVSALASNYWMLSLWTDPIVNGQDHTVRVLSVYALGISQIAVFGYSMA 961
QY 1037 VSIIGILASRCLHYDLHSLRSPMSFERTPSGNLVNRSKELDYDSMPEYIKMFMG 1096
DB 962 VSIIGILASRCLHYDLHSLRSPMSFERTPSGNLVNRSKELDYDSMPEYIKMFMG 1021
QY 1097 SLFNVIGACIVILLATPIAAIIPPLGLIYFVQRFVVASRQLKLESVRSRSPVYSHFN 1156
DB 1022 SLFNVIGACIVILLATPIAAIIPPLGLIYFVQRFVVASRQLKLESVRSRSPVYSHFN 1081
QY 1157 ETLGVSIVIRAFEBQERFIHQSDLKVDENQKAYYPSIVANRWLAVRLECVGNCIVLPAL 1216
DB 1082 ETLGVSIVIRAFEBQERFIHQSDLKVDENQKAYYPSIVANRWLAVRLECVGNCIVLPAL 1141
QY 1217 FAVISRHSLSAGLVGSLVSYSLQVTTYLNLVWSSSEMETNIVAEVLKEYSETEKAPW 1276
DB 1142 FAVISRHSLSAGLVGSLVSYSLQVTTYLNLVWSSSEMETNIVAEVLKEYSETEKAPW 1201
QY 1277 QIOETAPSSMPQGRFEPFNVCIRYREDLDPVLRHINVTINGEKYGIQRTAGKSSL 1336
DB 1202 QIOETAPSSMPQGRFEPFNVCIRYREDLDPVLRHINVTINGEKYGIQRTAGKSSL 1261
QY 1337 TLGFLRINSAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLFSSGLRMLDFFSOYSD 1396
DB 1262 TLGFLRINSAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLFSSGLRMLDFFSOYSD 1321
QY 1397 EEWTSLELAHLKDPVSALPDKLDHBCAEGENLSVGQROLVCLARALLRRTKILVIDEA 1456
DB 1322 EEWTSLELAHLKDPVSALPDKLDHBCAEGENLSVGQROLVCLARALLRRTKILVIDEA 1381
QY 1457 TAAVDLETDLLIOSTRFQEDCTVLTARLNTIMTYTVIVYDKIEQYGPSSLLQ 1516
DB 1382 TAAVDLETDLLIOSTRFQEDCTVLTARLNTIMTYTVIVYDKIEQYGPSSLLQ 1441
QY 1517 QRGFLYSMAKDAGLV 1531
DB 1442 QRGFLYSMAKDAGLV 1456

RESULT 7
O6UR05 CANPA
ID O6UR05 CANPA PRELIMINARY; PRT; 1531 AA.
AC O6UR05;
DT 05-JUL-2004 (Tremblere1. 27, Created)
DT 05-JUL-2004 (Tremblere1. 27, Last sequence update)
DT 05-JUL-2004 (Tremblere1. 27, Last annotation update)
DE Multidrug resistance-associated protein 1.
GN Name=MRP1;
OS Canis familiaris (dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
OC Canis.
OX NCBI_TaxID=9615;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22404678; Pubmed=12516967;
RA Ma L., Pratt S.B., Cao J., Danczig A.H., Moore R.E., Slapak C.A.;
RT "Identification and characterization of the canine multidrug
resistance-associated protein.";
RL Mol. Cancer Ther. 1:1335-1342(2002).
[2]

```

RP NUCLEOTIDE SEQUENCE.
Ma L., Pratt S.E., Cao J., Danczig A.H., Moore R.E., Slapak C.A.,
Submitted (Aug-2003) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL: AY63728; AAC3148.1; -; mRNA.
DR Ensembl: ENSCAFG00000018206; Canis familiaris.
DR GO: GO:0016021; C:Integral to membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0016887; F:ATPase activity; IEA.
DR GO: GO:0042656; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:0000165; F:nucleotide binding; IEA.
DR GO: GO:0006810; P:transporter; IEA.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR011527; ABC_membrane_1.
DR InterPro: IPR011440; ABC_TM_transp.
DR InterPro: IPR003439; ABC_transp_like.
DR InterPro: IPR005292; MRP_assoc.
DR InterPro: IPR000719; Proc_kinase.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD000006; ABC_transporter; 2.
DR SMART: SM00382; AAA; 2.
DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
DR PROSITE: PS50929; ABC_TMTP_2; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE: PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN 1.
KW ATP-binding; Nucleotide-binding; Repeat; Transport.
SQ SEQUENCE 1531 AA; 171792 MW; 746361A716158BD CRC64;

Query Match 73.4%; Score 7272; DB 2; Length 1531;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 1402; Conservative 70; Mismatches 59; Indels 0; Gaps 0;

QY 1 MALGFGCAGSDSDPLMDNMVMTNTSNPDTFCFQNTVAVWPCFYLMACFPYFLYSRHH 60
DB 1 MALGFGCAGSDSDPLMDNMVMTNTSNPDTFCFQNTVAVWPCFYLMACFPYFLYSRHH 60
QY 61 DRGYIOMTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPILLGITTLA 120
DB 61 DRGYIOMTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPILLGITTLA 120
QY 121 TFLIQLEERRKGVQSSGIMLTFMLVALVCALALRSKIMTALKEADAQVLDPRDITFYVFS 180
DB 121 TFLIQLEERRKGVQSSGIMLTFMLVALVCALALRSKIMTALKEADAQVLDPRDITFYVFS 180
QY 181 LLLIQVLVSCSDRSPLFSETHDNPCESSASFLSITFWMTIGLVRCGRQLBESD 240
DB 181 LLLIQVLVSCSDRSPLFSETHDNPCESSASFLSITFWMTIGLVRCGRQLBESD 240
QY 241 LMSLNKEDTSQOVVPLVKNMKKECAKTRKQPKVAVVSKDPAOPKRESSKYDANEVEAL 300
DB 241 LMSLNKEDTSQOVVPLVKNMKKECAKTRKQPKVAVVSKDPAOPKRESSKYDANEVEAL 300
QY 301 IVKSPQKEMWNSLFVLYKTFEGPYFLMGFFPKAHDLMFMFSGPQILKLLIFVNDYKAD 360
DB 301 IVKSPQKEMWNSLFVLYKTFEGPYFLMGFFPKAHDLMFMFSGPQILKLLIFVNDYKAD 360
QY 361 WOGFFYVLLFVTLAQTLVLAHQVFHICFVSGMRKTAIVIGAVYKAVITVNSAKKSTV 420
DB 361 WOGFFYVLLFVTLAQTLVLAHQVFHICFVSGMRKTAIVIGAVYKAVITVNSAKKSTV 420
QY 421 GEIYVIMSVDAQRFMDLATYINMINSAPLOVIALYLLMLNIGSPVLGAVAVMLVNVN 480
DB 421 GEIYVIMSVDAQRFMDLATYINMINSAPLOVIALYLLMLNIGSPVLGAVAVMLVNVN 480
QY 481 AVAMAKTKTYOVAAHKSKDNRIKLMNEILNGIKVLKLYAMELAFKQVLAIRQEBLKYLK 540
DB 481 AVAMAKTKTYOVAAHKSKDNRIKLMNEILNGIKVLKLYAMELAFKQVLAIRQEBLKYLK 540
QY 541 KSAIYLSANGTPTWCTPFLVALCTFAVYVTTDENNILLDAQTFVSLAFNIIREFLNTLP 600
DB 541 KSAIYLSANGTPTWCTPFLVALCTFAVYVTTDENNILLDAQTFVSLAFNIIREFLNTLP 600

QY 601 MVISSIVQASVSLKRLAIFLSHEELPDSIERRPVKDGGGNTSITVNAFTTWARSDPT 660
DB 601 MVISSIVQASVSLKRLAIFLSHEELPDSIERRPVKDGGGNTSITVNAFTTWARSDPT 660
QY 661 LMGITTSIPREALVAVVGVQCCGSSLLSALLAMDMVEGVAAIKGSVAVYPOQAMQND 720
DB 661 LMGITTSIPREALVAVVGVQCCGSSLLSALLAMDMVEGVAAIKGSVAVYPOQAMQND 720
QY 721 SLRNNILFGCOLLEPPYRSVIAQCALLPDEILPSGRTIEGKGVNLSGGOKORVSLAR 780
DB 721 SLRNNILFGCOLLEPPYRSVIAQCALLPDEILPSGRTIEGKGVNLSGGOKORVSLAR 780
QY 781 AVYSNADIIYFDDPLSAVDAAVHGHIIFENVIGPYGMKKNKTRILIVTSMSTLPQVDVIV 840
DB 781 AVYCDSDIYLPDDLSAVDAVHGHIIFENVIGPYGMKKNKTRILIVTSMSTLPQVDVIV 840
QY 841 MSGKISMGSYOELLARDDGFAFELTVAYSTBOEDABEAGVAVGSGPGKEAKOMENG 900
DB 841 MTGKISMGSYOELLARDDGFAFELTVAYSTBOEDABEAGVAVGSGPGKEAKOMENG 900
QY 901 LVYDSAGKOLQRLSSSSSYSGDISRRHNSFAELQKAKEETWTKLMEADKAQTGVYKL 960
DB 901 LVYDSAGKOLQRLSSSSSYSGDISRRHNSFAELQKAKEETWTKLMEADKAQTGVYKL 960
QY 961 SVYWDYMKALGLFISFLSIFLEMCNHVASALASNYWLSLWTDPIVNGTOEHTKYRLSYG 1020
DB 961 SVYWDYMKALGLFISFLSIFLEMCNHVASALASNYWLSLWTDPIVNGTOEHTKYRLSYG 1020
QY 1021 ALGISQGIYAVRGYMAVSIIGIILASRCHVDLHLSILRSPSPFERPISGULVNRFSKL 1080
DB 1021 ALGISQGIYAVRGYMAVSIIGIILASRCHVDLHLSILRSPSPFERPISGULVNRFSKL 1080
QY 1081 DTVDSMIPYVIMKMGSLFNIVIGACIVLLATPIAIIIPPLGIYFVQGFYVASSRQL 1140
DB 1081 DTVDSMIPYVIMKMGSLFNIVIGACIVLLATPIAIIIPPLGIYFVQGFYVASSRQL 1140
QY 1141 KRLESVSRSPYSHFNETLLGVSVYIRAFEEQERFIROSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLESVSRSPYSHFNETLLGVSVYIRAFEEQERFIROSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLCEVNGCIVLPAALPAVISRHSLSAGVLVSYSIQVTTYINMLVRMSSEMETNIVA 1260
DB 1201 VRLCEVNGCIVLPAALPAVISRHSLSAGVLVSYSIQVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VRLCEVNGCIVLPAALPAVISRHSLSAGVLVSYSIQVTTYINMLVRMSSEMETNIVA 1260
DB 1261 VRLCEVNGCIVLPAALPAVISRHSLSAGVLVSYSIQVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEVSETEKAPWQIOETAPSSWPQVGRVFRNYCLAREDLDFVLHINVTINGG 1320
DB 1261 VERLKEVSETEKAPWQIOETAPSSWPQVGRVFRNYCLAREDLDFVLHINVTINGG 1320
QY 1321 EKVGIVRTGAGKSLTGLFRINESAGEIILIDGINIAKIGLHDLRPKITIIPODPVLF 1380
DB 1321 EKVGIVRTGAGKSLTGLFRINESAGEIILIDGINIAKIGLHDLRPKITIIPODPVLF 1380
QY 1381 SGSLRPMNIDPFSQVSDDEVWTSLELAHKDVSAALPDLDBHECAEGENLSVGGQRLVCL 1440
DB 1381 SGSLRPMNIDPFSQVSDDEVWTSLELAHKDVSAALPDLDBHECAEGENLSVGGQRLVCL 1440
QY 1441 ARALLRKTIIIVLEBATAVDELTDLLQSTIRPQEDCTVLTJAHRANTIMDYTRYVIL 1500
DB 1441 ARALLRKTIIIVLEBATAVDELTDLLQSTIRPQEDCTVLTJAHRANTIMDYTRYVIL 1500
QY 1501 DKGEIOEYGAPSDILLQGRGLFYSSAKDAGLV 1531
DB 1501 DKGEIOEYGAPSDILLQGRGLFYSSAKDAGLV 1531

RESULT 8
059G19 HUMAN
ID 059G19 HUMAN PRELIMINARY; PRT; 1439 AA.
AC 059G19
DT 10-MAY-2005 (Tremblrel. 30, Created)
DT 10-MAY-2005 (Tremblrel. 30, Last sequence update)
DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)

DE ATP-binding cassette, sub-family C, member 1 isoform 1 variant
(Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain.
RA TOOKA Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.,
RT "None Title."
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SMITABILITY: Belongs to the ABC transporter family.
DR EMBL; AB209120; BAD92357.1; -; mRNA.
DR Ensemble; ENSG00000103222; Homo sapiens.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC membrane 1.
DR InterPro; IPR001140; ABC_TM_transport.
DR InterPro; IPR003439; ABC_TM_transport.
DR InterPro; IPR005292; MRP_assoc.
DR InterPro; IPR007019; Prot_kinase.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN KINASE_ATP; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding.
FT NON_TER 1
SQ SEQUENCE 1439 AA; 160439 MW; 090D881ADBFB3077 CRC64;
Query Match 73.1%; Score 7242; DB 2; Length 1439;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1428; Conservative 0; Mismatches 1; Indels 10; Gaps 1;
QY 103 APVFLVSPPTLIGITTLTATFLIQLERRKGVSGLMTFMLVALVCAIILRSKIMTALK 162
DB 1 APVFLVSPPTLIGITMLATFLIQLERRKGVSGLMTFMLVALVCAIILRSKIMTALK 60
QY 163 EDQAVDLPDITFVYFSSLLIQVLVSCFSDRSLFSETIHDPRCPRESSASFLSRTTFW 222
DB 61 EDQAVDLPDITFVYFSSLLIQVLVSCFSDRSLFSETIHDPRCPRESSASFLSRTTFW 120
QY 223 WITGLIYRGYROPLEGSDLSINKEDTSEGVPLVYVNMWKECAKTRKOPKYVYSSKDP 282
DB 121 WITGLIYRGYROPLEGSDLSINKEDTSEGVPLVYVNMWKECAKTRKOPKYVYSSKDP 180
QY 283 AQPRESSKVDANEVEALIVKSPQKWNPSLFEVLYTFFGYFLMSFFFAIHDLMFSG 342
DB 181 AQPRESSKVDANEVEALIVKSPQKWNPSLFEVLYTFFGYFLMSFFFAIHDLMFSG 240
QY 343 POLIKLLIKFVNPTKADNOCGYFTYVLLFTYACIQTLVLAQYFHICVSGMRITAVIGA 402
DB 241 POLIKLLIKFVNPTKADNOCGYFTYVLLFTYACIQTLVLAQYFHICVSGMRITAVIGA 300
QY 403 VYRKALVITNSARKSSVGEIIVNLSYDAQRFMDLATYINMWSAPLOVILLATVLMNL 462
DB 301 VYRKALVITNSARKSSVGEIIVNLSYDAQRFMDLATYINMWSAPLOVILLATVLMNL 360
QY 463 GPSTLAVAVMTLVPNVAVAMAKTKTYOVAHMSKNRIKLNELINGIKVLKYLAMEL 522
DB 361 GPSTLAVAVMTLVPNVAVAMAKTKTYOVAHMSKNRIKLNELINGIKVLKYLAMEL 420

QY 523 AFKDKVLAIROBELKVLKKSAYLSAVGTFTWCTPFLVALCTFAVYVYIDENNILDAQTA 582
DB 421 AFKDKVLAIROBELKVLKKSAYLSAVGTFTWCTPFLVALCTFAVYVYIDENNILDAQTA 480
QY 583 FVSLALFNILRFPNLILPMVYSSIVQASVSLKRLRIFLSHELEPDSIERRPVQGGGTN 642
DB 481 FVSLALFNILRFPNLILPMVYSSIVQASVSLKRLRIFLSHELEPDSIERRPVQGGGTN 540
QY 643 SITVRNATFTWASDPPTLNGITFSPLEGALVAVVGVGGKSSLSALLAENDKVEGHV 702
DB 541 SITVRNATFTWASDPPTLNGITFSPLEGALVAVVGVGGKSSLSALLAENDKVEGHV 600
QY 703 AIKGSVAVVQOAMVQWDSIRENILEFGCQLEEPYRSVIOACALLPDLTLPSCDREIG 762
DB 601 AIKGSVAVVQOAMVQWDSIRENILEFGCQLEEPYRSVIOACALLPDLTLPSCDREIG 660
QY 763 EKVNLISGQKORVSLARAVYSNADITLFPDPLSAVDAAHVKHIFENVIGPKMLKNKTR 822
DB 661 EKVNLISGQKORVSLARAVYSNADITLFPDPLSAVDAAHVKHIFENVIGPKMLKNKTR 720
QY 823 ILVTHSMSTYLPQVDVITVMSGGKISEMGSYQELIARDGAFAPLRTYASTBOBDAEEN 881
DB 721 ILVTHSMSTYLPQVDVITVMSGGKISEMGSYQELIARDGAFAPLRTYASTBOBDAEEN 780
QY 882 -----GVTVSGPGKAKOMENGMLVTDASAGKQLORQLSSSSSGSDISRHNSTA 932
DB 781 STWDEEEAAGVTVSGGKAKOMENGMLVTDASAGKQLORQLSSSSSGSDISRHNSTA 840
QY 933 ELQKAEKKEETKWLMEADRAQTQVGLSYWMDYMKAGHIFISFLSIFLMCHVSLAS 992
DB 841 ELQKAEKKEETKWLMEADRAQTQVGLSYWMDYMKAGHIFISFLSIFLMCHVSLAS 900
QY 993 NYMLSLMTDDPIYNGTOEHTKVLRYGALGISQGIARFVGYSAVSGIILASRCHVDL 1052
DB 901 NYMLSLMTDDPIYNGTOEHTKVLRYGALGISQGIARFVGYSAVSGIILASRCHVDL 960
QY 1053 LHSILSPMSFPERTSQGNLYNRFSKELDVDSMIEBVIKMFNGSLFNVIQACIVILLAT 1112
DB 961 LHSILSPMSFPERTSQGNLYNRFSKELDVDSMIEBVIKMFNGSLFNVIQACIVILLAT 1020
QY 1113 PIAAIIIPPLGLIYFFQRFYVASSRQLKRLSESVRSPPVSHFNETLLAGSVIRAFEEQ 1172
DB 1021 PIAAIIIPPLGLIYFFQRFYVASSRQLKRLSESVRSPPVSHFNETLLAGSVIRAFEEQ 1080
QY 1173 RFIHQSDLKVDENQAKAYPSIVANRWLAVERLECVGNCIVLPAALPAVISRHSLSAGLV 1232
DB 1081 RFIHQSDLKVDENQAKAYPSIVANRWLAVERLECVGNCIVLPAALPAVISRHSLSAGLV 1140
QY 1233 SVSISLQVTTYLNMVLRMSSEMETNIYAVERLKEYSETEKEAPWQOETAPESSWPQVR 1292
DB 1141 SVSISLQVTTYLNMVLRMSSEMETNIYAVERLKEYSETEKEAPWQOETAPESSWPQVR 1200
QY 1293 VEFRRNYCLARYEDLDVLRHINVTINGEKVGIYGRGTGAKSLSLTGLPRINSAGEII 1352
DB 1201 VEFRRNYCLARYEDLDVLRHINVTINGEKVGIYGRGTGAKSLSLTGLPRINSAGEII 1260
QY 1353 IDGINIAKIGLHDLRFKRTITIIPODPVLFSGSLRMNLDPSQYSDEEWTSLSLAHLKDFV 1412
DB 1261 IDGINIAKIGLHDLRFKRTITIIPODPVLFSGSLRMNLDPSQYSDEEWTSLSLAHLKDFV 1320
QY 1413 SALPDKLDHEHCAGGENLSVGROQLVCLARALLRKRLIIVLDATAVAVDLETDDLQSTI 1472
DB 1321 SALPDKLDHEHCAGGENLSVGROQLVCLARALLRKRLIIVLDATAVAVDLETDDLQSTI 1380
QY 1473 RTQFEDCTVLTIAHRLNTIMDYTRVIVLQKGEIQEAGPSDDLQOQGLFYSMAKDAGLV 1531
DB 1381 RTQFEDCTVLTIAHRLNTIMDYTRVIVLQKGEIQEAGPSDDLQOQGLFYSMAKDAGLV 1439
RESULT 9
Q8HX05_BOVIN
ID Q8HX05_BOVIN PRELIMINARY; PRT; 1530 AA.
AC Q8HX05;

DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Multidrug resistance protein 1.
GN Name=MRP1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22062955; PubMed=12067707; DOI=10.1016/S0014-5793(02)02816-8;
RA Taguchi T., Saeki K., Komano T.;
RT "Functional analysis of MRP1 cloned from bovine."
RL FEBS Lett. 521:211-213(2002).
CC -1-SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AB082124; BAC15550.1; -; mRNA.
DR HSSP; P08716; IMTO.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_Atpase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMIF_2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR ATP-binding; Nucleotide-binding; Repeat.
KW SEQUENCE 1530 AA; 171666 MW; AABAF92ED7832703 CRC64;
SQ
Query Match 73.0%; Score 7230.5; DB 2; Length 1530;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1388; Conservative 83; Mismatches 59; Indels 1; Gaps 1;
QY 1 MALRGFCAGDSDPFLMDNMVNTNTSPDFTKCFQNTVAVWVCFYLMAFCPPYFLYLSRH 60
DB 1 MALRDFGVSDSDFLFEMNNTVNTNSPDFTKCFQNTVAVWVCFYLMAFCPPYFLYLSRH 60
QY 61 DRGYIQMTPLNKTATLGFLLMIYVCMADLFYSFWERSRGIFLAPVFLVSPILLGITLLA 120
DB 61 DRGYIQMTHLAKATLGFLLMIYVCMADLFYSFWERSMGKLLAPVFLVSPILLGITLLA 120
QY 121 TFLIQLERKGVQSSGIMLTFWVALVCAALALRSKIMTALKEDQVDLFRDITRYVFS 180
DB 121 TFLIQLERRRGVQSSGIMLTFWLIALCALILRSKIMTALKEDARVAVFRDITRYVFS 180
QY 121 TFLIQLERRRGVQSSGIMLTFWLIALCALILRSKIMTALKEDARVAVFRDITRYVFS 180
DB 121 TFLIQLERRRGVQSSGIMLTFWLIALCALILRSKIMTALKEDARVAVFRDITRYVFS 180
QY 181 LLLIQLVSCFSDRSLPSETIHDPNCPESASFLSITTFWMIINGLIVRGROPLEGSD 240
DB 181 LVLILQLVSCFSDRSLPSETIHDPNCPESASFLSITTFWMIINGLIVRGROPLEGSD 240
QY 241 LMSLNKEDTSEQVAVLVKNNKKECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVAL 300
DB 241 LMSLNKEDTSEQVAVLVKNNKKECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVAL 300
QY 301 IVKSPQKEMNPSLFVLYKTEGPIYVLMSEFFKAIHDLMMFSGPOLIKLIFVNDTKAPD 360
DB 301 IVKSPQKEMNPSLFVLYKTEGPIYVLMSEFFKAIHDLMMFSGPOLIKLIFVNDTKAPD 360
QY 361 WQGFYTYTLVLTFLAQLQTLVLAHQVHICFVSGMRIKTAVIGAVYKKAIVTMSAKKSSIV 420
DB 361 WQGFYTYTLVLTFLAQLQTLVLAHQVHICFVSGMRIKTAVIGAVYKKAIVTMSAKKSSIV 420

QY 421 GEIVNLSVDAQRMEDATTINMTWSAPLOYITLALYLMLNLGSSVLAGVAVNMLAMPVN 480
DB 421 GEIVNLSVDAQRMEDATTINMTWSAPLOYITLALYLMLNLGSSVLAGVAVNMLAMPVN 480
QY 481 AVMAKTKTQVAMKSKDNRIKLMNEILNGIKYLKLYAMELAKDKVLAIRQELKYLK 540
DB 481 AVMAKTKTQVAMKSKDNRIKLMNEILNGIKYLKLYAMELAKDKVLAIRQELKYLK 540
QY 541 KSAVLSAVGFTWCTPFLVALCTFAVYVITIDENNIIIDAQTAFLVALFNILREPLNLP 600
DB 541 KSAVLSAVGFTWCTPFLVALCTFAVYVITIDENNIIIDAQTAFLVALFNILREPLNLP 600
QY 601 MVISIYQASVSLKRLAIFLSHSELDPDSIFERRPVKQGGGNTSTVKNATFTWASDPPT 660
DB 601 MVISIYQASVSLKRLAIFLSHSELDPDSIFERRPVKQGGGNTSTVKNATFTWASDPPT 660
QY 661 LMGITFSPGALVAVVGVGCGKSSLLSLALAMDYEGHVALKGSAAVYPOQAMIOND 720
DB 661 LMGITFSPGALVAVVGVGCGKSSLLSLALAMDYEGHVALKGSAAVYPOQAMIOND 720
QY 721 SLRENILFGCOLBEPYRSVLIQACALLPDEILPSGDRTEIGEGVNLGGQKQKRVSLAR 780
DB 721 SLRENILFGQRLQGRYYKAVVEACALLPDEILPSGDRTEIGEGVNLGGQKQKRVSLAR 780
QY 781 AVYSNADITYLPDDL SAVDAVGHGIFBNYIGPFGMLKNTKRLIVTMSGLPQVDYIV 840
DB 781 AVYCDSDVYLLDDPLSAVDVGHGIFBNYIGPFGMLKNTKRLIVTMSGLPQVDYIV 840
QY 841 MSGGKISPMGSYQELLARDAFAEFLLRTYASTBEOBDAEENGVYSGPGKEAOMENGM 900
DB 841 MSGGKISPMGSYQELLARDAFAEFLLRTYASTBEOBDAEENGVYSGPGKEAOMENGM 900
QY 901 LVTSAGKOLQRLQSSSSSYSGDISRHNSYAELOKA BAKKEBTWKLMEADKAQTQVYL 960
DB 901 LVTSAGKOLQRLQSSSSSYSGDISRHNSYAELOKA BAKKEBTWKLMEADKAQTQVYL 960
QY 961 SVYWDYKAIQLFISFISIFLPMGNHVSALASNYWLSLWTDLPVNGTOEHTKRLSYG 1020
DB 961 SVYWDYKAIQLFISFISIFLPMGNHVSALASNYWLSLWTDLPVNGTOEHTKRLSYG 1020
QY 1021 ALGISOGIAVFGVMAVSIIGIILASRCILHDLHSIIRSPSPFERPISGLVNRFSKEL 1080
DB 1021 ALGISOGIAVFGVMAVSIIGIILASRCILHDLHSIIRSPSPFERPISGLVNRFSKEL 1080
QY 1080 DTVDSMIPQVIMKMGSLFNVIAGCIYLLATPIAIIIPPLGIYFFVGQFYVAASSROL 1140
DB 1080 DTVDSMIPQVIMKMGSLFNVIAGCIYLLATPIAIIIPPLGIYFFVGQFYVAASSROL 1140
QY 1141 KRLSVSRSPVYSHPNETLLGVSVITRAFEQERFIHQSDLKVDENOKAYYSIVANRWLA 1200
DB 1141 KRLSVSRSPVYSHPNETLLGVSVITRAFEQERFIHQSDLKVDENOKAYYSIVANRWLA 1200
QY 1201 VRLCEVNGCIYLPALPAVVISRHSLSAGIVGLSVSYSLQVTTYVNLVVRMSEMETNIVA 1260
DB 1201 VRLCEVNGCIYLPALPAVVISRHSLSAGIVGLSVSYSLQVTTYVNLVVRMSEMETNIVA 1260
QY 1261 VERLKEVSETEKAPMOIQETAPPSWPOVQVRFVFRNYCLTYREDDLDFVLHINVTING 1320
DB 1261 VERLKEVSETEKAPMOIQETAPPSWPOVQVRFVFRNYCLTYREDDLDFVLHINVTING 1320
QY 1320 EKVGVIGRTGAKSSLTGLFRIKESAGEI IIDDINIAKIGLHDLRFKTIITIPQDPLF 1379
DB 1320 EKVGVIGRTGAKSSLTGLFRIKESAGEI IIDDINIAKIGLHDLRFKTIITIPQDPLF 1379
QY 1381 SGSLRPNLDPFSQYSDERWTSLELAHKDVSALPKDLBHCAGGENLSVGOQOLVCL 1440
DB 1381 SGSLRPNLDPFSQYSDERWTSLELAHKDVSALPKDLBHCAGGENLSVGOQOLVCL 1440
QY 1441 ABALLRKTILVLEBATAVVDLETDLIQSTIRIROPEDCYLTTIAHRLNTIMDTTRVIVL 1500
DB 1441 ABALLRKTILVLEBATAVVDLETDLIQSTIRIROPEDCYLTTIAHRLNTIMDTTRVIVL 1500
QY 1501 DKGEIOEYGA PSDLLQQRGLFY SMAKDAGLV 1531
DB 1501 DKGEIOEYGA PSDLLQQRGLFY SMAKDAGLV 1531

QY 977 LSTFLMNCNHSALASNTWLSMTDDPLVNGTOETKRLSYGALGISOGIANFGYSMA 1036
 DB 846 LSTFLMNCNHSALASNTWLSMTDDPLVNGTOETKRLSYGALGISOGIANFGYSMA 905
 QY 1037 VSTIGIIASRCIHDLSIIIRSPMSFPERIPSGNLVNRFSKEDTVDMSIPEVIXMGMG 1096
 DB 906 VSTIGIIASRCIHDLSIIIRSPMSFPERIPSGNLVNRFSKEDTVDMSIPEVIXMGMG 965
 QY 1097 SLFNVIGACIYILLATPIAIIIPPLGLIYFVQRFYVASSRQRLKRLSVSPYSHFN 1156
 DB 966 SLFNVIGACIYILLATPIAIIIPPLGLIYFVQRFYVASSRQRLKRLSVSPYSHFN 1025
 QY 1157 ETLIGVSTIRAFEEDEERFIHQSDLKVDENOKAYYSIYANRWLANRLECVGNCIYLFAL 1216
 DB 1026 ETLIGVSTIRAFEEDEERFIHQSDLKVDENOKAYYSIYANRWLANRLECVGNCIYLFAL 1085
 QY 1217 FAVISRHSLSAGLVGSYSIQVTTYANMLVRMSSEMETNIIVAVERLKEYSETEKAPW 1276
 DB 1086 FAVISRHSLSAGLVGSYSIQVTTYANMLVRMSSEMETNIIVAVERLKEYSETEKAPW 1145
 QY 1277 QIOETAPPSWPQVGRVFRNVCLEFRREDLPVLRIHINVTINGEKVGIWRTGAGKSSL 1336
 DB 1146 QIOETAPPSWPQVGRVFRNVCLEFRREDLPVLRIHINVTINGEKVGIWRTGAGKSSL 1205
 QY 1337 TLGLFRINESABGEIITIDGINIAKIGLHDLRFKTIITIPQDPVLFSGSLRMLDPPSQYSD 1396
 DB 1206 TLGLFRINESABGEIITIDGINIAKIGLHDLRFKTIITIPQDPVLFSGSLRMLDPPSQYSD 1265
 QY 1397 BEWMTSLFLAHKDPVSALPDGLDHECAEGGENLSVSGROLYCLARALLRKTKILVLEBA 1456
 DB 1266 BEWMTSLFLAHKDPVSALPDGLDHECAEGGENLSVSGROLYCLARALLRKTKILVLEBA 1325
 QY 1457 TAAVDLETDLIQSTIRTOFEDCTVLTAAHRLNTIMDYTRVILDKGEIOEGAPSDLIQ 1516
 DB 1326 TAAVDLETDLIQSTIRTOFEDCTVLTAAHRLNTIMDYTRVILDKGEIOEGAPSDLIQ 1385
 QY 1517 QRGLEFYSMAKDAGLV 1531
 DB 1386 QRGLEFYSMAKDAGLV 1400

RESULT 11
 MRP1_MOUSE STANDARD; PRT; 1528 AA.
 AC 035379;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Multidrug resistance-associated protein 1 (ATP-binding cassette sub-family C member 1).
 GN Name=Abcc1; Synonyms=Abcc1a, Abcc1b, Mdrap, Mrp;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; OC Muroidea; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RX NUCLEOTIDE SEQUENCE.
 RC TISSUE=Muscle.
 RX MEDLINE=96251691; PubMed=8649356;
 RA Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P.C., Deeley R.G.;
 RT "Structure and expression of the messenger RNA encoding the murine multidrug resistance protein, an ATP-binding cassette transporter.",
 RL Mol. Pharmacol. 49:962-971(1996).
 RN (2)
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=22354683; PubMed=1246851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Nikiido I., Otsu N., Saito R., Suzuki H., Yamataka I., Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T., Baldarelli R., Hall D.P., Bult C., Hume D.A., Quackenbush J.,

RA Schrieml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chochia C., Cordani L.E., Cousins S., Dala B., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S., RA Gaesteland T., Gariboldi M., Giesi C., Godzik A., Gough J., RA Giromond S., Guetlich S., Hirokawa N., Jackson I.J., Jarvis E.D., RA Kanai A., Kawaji H., Kawasawa Y., Kedzierzki R.M., King B.L., RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A., RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H., RA Nagasima T., Numa K., Okido T., Pavan W.J., Petrea G., Pesole G., RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M., RA Sadelin A., Schneider C., Sempile C.A., Setou M., Shimada K., RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M., RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C., RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L., RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., RA Hitzane-Kishikawa T., Kono H., Nakamura M., Sakazune N., Sato K., RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S., RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I., RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa J., RA Yaenushli A., Yoshino M., Waterston R., Lander E.S., Rogers J., RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 CC -1- FUNCTION: May participate directly in the active transport of drugs into subcellular organelles or influence drug distribution indirectly (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.
 CC -1- SIMILARITY: Contains 2 ABC transmembrane type-1 domains.
 CC -1- SIMILARITY: Contains 2 ABC transporter domains.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
 CC -----
 CC EMBL; AF022908; AAB80938.1; -; mRNA.
 CC EMBL; AK029876; BAC26654.1; -; mRNA.
 CC HSSP; P08716; 1MT0.
 CC ENSEMBL; ENSMUSG00000023088; Mus musculus.
 DR MGI; MGI:102676; Abcc1.
 DR GO; GO:0005887; C:integral to plasma membrane; IDA.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR011527; ABC_membrane_1.
 DR InterPro; IPR001140; ABC_TM_transp.
 DR InterPro; IPR003439; ABC_transp_like.
 DR InterPro; IPR005292; MRP_assoc.
 DR Pfam; PF00664; ABC_membrane; 2.
 DR Pfam; PF00005; ABC_tran; 2.
 DR ProDom; PD000006; ABC_transporter; 2.
 DR SMART; SM00382; AAA; 2.
 DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
 DR PROSITE; PS50929; ABC_TM1F; 2.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
 DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
 KW ATP-binding; Glycoprotein; Nucleotide-binding; Repeat; Transmembrane; Transport.
 KM TOPO_DOM 1 33 Extracellular (By similarity).
 FT TOPO_DOM 34 54 1 (By similarity).
 FT TOPO_DOM 55 74 Cytoplasmic (By similarity).
 FT TRANSMEM 75 95 2 (By similarity).
 FT TOPO_DOM 96 100 Extracellular (By similarity).
 FT TRANSMEM 101 121 3 (By similarity).
 FT TOPO_DOM 122 133 Cytoplasmic (By similarity).
 FT TRANSMEM 134 154 4 (By similarity).
 FT TOPO_DOM 155 172 Extracellular (By similarity).
 FT TRANSMEM 173 193 5 (By similarity).
 FT TOPO_DOM 194 317 Cytoplasmic (By similarity).
 FT TRANSMEM 318 338 6 (By similarity).
 FT TOPO_DOM 339 364 Extracellular (By similarity).

FT	TRANSMEM	365	385	7 (By similarity).
FT	TOPO_DOM	386	441	Cytoplasmic (By similarity).
FT	TRANSMEM	442	462	8 (By similarity).
FT	TOPO_DOM	463	465	Extracellular (By similarity).
FT	TRANSMEM	466	486	9 (By similarity).
FT	TOPO_DOM	487	548	Cytoplasmic (By similarity).
FT	TRANSMEM	549	569	10 (By similarity).
FT	TOPO_DOM	570	591	Extracellular (By similarity).
FT	TRANSMEM	592	612	11 (By similarity).
FT	TOPO_DOM	613	963	Cytoplasmic (By similarity).
FT	TRANSMEM	964	984	12 (By similarity).
FT	TOPO_DOM	985	1022	Extracellular (By similarity).
FT	TRANSMEM	1023	1043	13 (By similarity).
FT	TOPO_DOM	1044	1086	Cytoplasmic (By similarity).
FT	TRANSMEM	1087	1107	14 (By similarity).
FT	TOPO_DOM	1108	1108	Extracellular (By similarity).
FT	TRANSMEM	1109	1129	15 (By similarity).
FT	TOPO_DOM	1130	1200	Cytoplasmic (By similarity).
FT	TRANSMEM	1201	1221	16 (By similarity).
FT	TOPO_DOM	1222	1223	Extracellular (By similarity).
FT	TRANSMEM	1224	1244	17 (By similarity).
FT	TOPO_DOM	1245	1528	Cytoplasmic (By similarity).
FT	DOMAIN	326	609	ABC transmembrane type-1 1.
FT	DOMAIN	644	868	ABC transporter 1.
FT	DOMAIN	971	1253	ABC transmembrane type-1 2.
FT	DOMAIN	1290	1524	ABC transporter 2.
FT	NP_BIND	678	685	ATP 1 (potential).
FT	NP_BIND	1324	1331	ATP 2 (potential).
FT	CARBOHYD	19	19	N-linked (GlcNAc...) (potential).
FT	CARBOHYD	1003	1003	N-linked (GlcNAc...) (potential).
EQ	SEQUENCE	1528 AA,	17185 MM,	68FD13667D61DB8 CRC64;

	Query Match	Best Local Similarity	70.7%;	Score 7002.5;	DB 1;	Length 1528;
	Matches 1349;	Conservative 102;	Mismatches 75;	Indels 7;	Gaps 5;	
QY	1	MALRGFCSADGSDPLMDMNVMTWNTSNPDFTKCFONTLVWVCFEYIMACPFPEFYLSRH	60			
Db	1	MALRSFCSADGSDPLMDMNVMTWNTSNPDFTKCFONTLVWVCFEYIMACPFPEFYLSRH	60			
QY	61	DRGITQMTPLNKTITLALGFLMIYCMADLFPSFMRSGITFLAPYFLVSPFLGLITLLA	120			
Db	61	DRGITQMTPLNKTITLALGFLMIYCMADLFPSFMRSGITFLAPYFLVSPFLGLITLLA	120			
QY	121	TFLLQLEBRKVOSSGIMLTFMVLAVCALIILSKITMTALKEPDQVDFPDIFPYYFS	180			
Db	121	TFLLQLEBRKVOSSGIMLTFMVLAVCALIILSKITMTALKEPDQVDFPDIFPYYFS	180			
QY	181	LLLIQVLVSCFSDRSPLESETIHDPNCPRESSASFLRITFWITGLIVRGYROPLEGSD	240			
Db	181	LLVLQVLVSCFSDRSPLESETIHDPNCPRESSASFLRITFWITGLIWMVHYRQPLESSD	240			
QY	241	LMSLNKEDTSQVYPVLVKMKKSCAKTRKQPPVYYS-SKDPAPKSSRYDANEVEYA	299			
Db	241	LMSLNKEDTSBEVYPVLVNNMKKSCCDKSRKQPPVIVAPADPSPKPKSSOLDVANEVEYA	300			
QY	300	LIIVSPKEMNPSPFLKVLKTFEGGYFLMSFFPKAIHDIIMMSGQIILKLLKFPNDITAP	359			
Db	301	LIIVSPKDRPSPFLKVLKTFEGGYFLMSFLYKALHDIMMAGPFIETLILINPNDREAP	360			
QY	360	DMQGYFYTVLLFTVACLQTLVLAHQYFHCIFVSGNRIKTAVIGAYRRKALVITNSARKSST	419			
Db	361	DMQGYFYTALLFTVSAQLQTLALHQYFHCIFVSGNRIKTAVIGAYRRKALLITNARKSST	420			
QY	420	VGEIVINMSVDAQRFMDLATYINNTMSAPLOVITLALYLLMLNIGPSVLAGYANVWLMPV	479			
Db	421	VGEIVINMSVDAQRFMDLATYINNTMSAPLOVITLALYLLMLNIGPSVLAGYANVWLMPV	480			
QY	480	NAVAMTKTKYQVAHMKSKDNRIKLMLNIGIKVLKLYAMELAFKQDYVILAROBELKVL	539			
Db	481	NAVAMTKTKYQVAHMKSKDNRIKLMLNIGIKVLKLYAMELAFQDKYMSIRQBELKVL	540			
QY	540	KKSAIVLSAVGTFTVCTPELVALCTPAVYVTIDENNIIADQATVSLAFNLIRPLPIL	599			

Db	541	KKSAIYLA	AVGTFTWCTCPFLVALSTFAVFTVBERNLIDAKAFVSLAFENILEFPLNLT	600
Qy	600	PMVTSIVQASVSLKRLRIFLSHLEELPDSIERPVDGGGTSITVRNAFTFARSDDP	659	
Db	601	PMVTSIVQASVSLKRLRIFLSHLEELPDSIERISISGSG-NSITYKMTFTAREBPP	659	
Qy	660	TLNGITFSIPGALVAVVGGVGGKSSLSALLAEMDKVEGHVAIKGSVAVVPOQAMION	719	
Db	660	TLNGITFSIPGALVAVVGGVGGKSSLSALLAEMDKVEGHVTLKGSVAVVPOQAMION	719	
Qy	720	DSJRENTLPECCQAEPPYRBSVIOACALLPLELILPBGDRTEIGGKGNLSGGCKORVSLA	779	
Db	720	DSJRENTLPEGHPIQENYIKKAWMEKCALLPLELILPBGDRTEIGKGNLSGGCKORVSLA	779	
Qy	780	RAYSNADITLFPDPLSAVDAAHAKGHIFENVYIGKGMKKKTRILVTHSMSYTPQVYVII	839	
Db	780	RAYSNADITLFPDPLSAVDAAHAKGHIFEKVGVGMGLKKKTRILVTHGISTYLPQVYVII	839	
Qy	840	VMSGGKISEMGSIYOELLARDGAPAEFLRTYASTEQEOBDAEENGTVGVSQPKKAKOMENG	899	
Db	840	VMSGGKISEMGSIYOELLARDGAPAEFLRTYANAEOQLASBDD--SVSGGKESKSPVENG	896	
Qy	900	MLVTDVQKQLOQOLSSSSSYSGDISKHNHSTAELOAEAKKEETWLMEMADKAQTOQYK	959	
Db	897	MLVTDVQKQLOQOLSSSSSYSGDISKHNHSTAELOAEAKKEETWLMEMADKAQTOQVQ	955	
Qy	960	LSVYMDPKAIGHIFISFLSIFLPMCNHVSALASVYMLSTWTD--PIYNGTOEHRKAVLSV	1018	
Db	956	LSVYMDPKAIGHIFITPLSIFLELCNHVSALASVYMLSTWTDPEVNGTOANNEFLSV	1015	
Qy	1019	YGALGISQGIATVFGYSMAVSIIGILASRCHAVDLHSLRSFMSFPERTPSGMLVNNFSK	1078	
Db	1016	YGALGISQGAATFGYSMAVSIIGIFASRRHLHDLVNVLASFMSFPERTPSGMLVNNFSK	1075	
Qy	1079	ELDTVDSMIREPVYIKMFNGSLFNNYIGACTIVLTLATPIAIIIPPLGLIYFVQRFYVASSR	1138	
Db	1076	ELDTVDSMIREPVYIKMFNGSLFVSYGAVIITLALTPIAVVIIPPLGLIYFVQRFYVASSR	1135	
Qy	1139	QLKRLSESVSRSPPYSHSHNETLLGVSYVRAFEBOBRFHOSDLKVDENQKAYPSBITVANRW	1198	
Db	1136	QLKRLSESVSRSPPYSHSHNETLLGVSYVRAFEBOBRFHOSDLKVDENQKAYPSBITVANRW	1195	
Qy	1199	LAVRLEECVGNCIYLFALPAVISRHSLSAGLVGLSVSYSLOVTVYANMLVRMSEMEETNI	1258	
Db	1196	LAVRLEECVGNCIYLFALPAVISRHSLSAGLVGLSVSYSLOVTVYANMLVRMSEMEETNI	1255	
Qy	1259	VAVRLEKESYETERKAPWQIOETAPRPSWPQVGRVEFRNYCLARYEDLDVFLRHINVTIN	1318	
Db	1256	VAVRLEKESYETERKAPWQIOETAPRPSWHSGRVEFRDYCLARYEDLDVFLKHNVTIE	1315	
Qy	1319	GGKRVGVGTGAGKSSLTGLRINESAGELIIIDGINIAKIGLHDLRKRITIIIPDDV	1378	
Db	1316	GGKRVGVGTGAGKSSLTGLRINESAGELIIDGVNIAKIGLHDLRKRITIIIPDDV	1375	
Qy	1379	LFGSGLRMMNDPFSQYSDDEVMTSLSLAHKDDPVSAIPDLTDBHECAGGENTL SVGOROLV	1438	
Db	1376	LFGSGLRMMNDPFSQYSDDEVMTSLSLAHKGVSAIPDLTDBHECAGGENTL SVGOROLV	1435	
Qy	1439	CLAPALIRKTXIILVDEATAVADLETFDDLIQSTIRFOEDCTVLTIAHRLNTIMDYTRVI	1498	
Db	1436	CLAPALIRKTXIILVDEATAVADLETFDDLIQSTIRFOEDCTVLTIAHRLNTIMDYTRVI	1495	
Qy	1499	VLDKGEIOEYGAPSDILQORGLFYMAKQAGLV 1531		
Db	1496	VLDKGEYRECGAPSELLOQRGIFYMAKQAGLV 1528		
RESULT 12				
Q810E4 RAT PRELIMINARY; PRT; 1532 AA.				
AC Q810E4; Q810E4; 01-JUN-2003 (TTEMBirel. 24, Created)				

DT 01-JUN-2003 (TRENBLREL. 24, last sequence update)
DT 01-MAR-2004 (TRENBLREL. 26, last annotation update)
DR ATP-binding cassette protein C1.
GN Name=Abcc1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Eularchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley, TISSUE=Spleen;
RA Yabuuchi H., Takayanagi S., Ishikawa T.;
RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF487549; AA085437.1; -, mRNA.
DR HSPSP; P08716; IMT0.
DR Ensembl; ENSRNOG0000032748; Rattus norvegicus.
DR RGD; 3112; Abcc1.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0006887; F:ATPase activity; IEA.
DR GO; GO:0042826; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_ATPase.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR001395; Aldo/kec_red.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00063; ALDOKEETO_REDUCTASE_3; UNKNOWN_1.
KM ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1532 AA; 171493 MW; 2E6939F63F5A3F68 CRC64;
Query Match 70.0%; Score 6932.5; DB 2; Length 1532;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 1339; Conservative 99; Mismatches 92; Indels 3; Gaps 3;
QY 1 MALRGFCGADGSDPLMDNVTWNTSNPDFTKCFQNTVLVWVPCFYLMACFPYFLYLSRH 60
DB 1 MALSSFCSSDSDPLMDNVTWNTSNPDFTKCFQNTVLVWVPCFYLMSCFPYFLYLSRH 60
QY 61 DRGYQMTPLNKTAKALGFLIMIVGADLFYSFMRSGIPLAPVPLVSPITLITTLA 120
DB 61 DRGYQMTPLNKTAKALGFLIMIVGADLFYSFMRSGIPLAPVPLVSPITLITTLA 120
QY 121 TFLQLERKGVSSGIMLTFWLVALVCALATLRKIMTALKEHQVDFRDIIFYFYS 180
DB 121 TFLQLERKGVSSGIMLTFWLVALVCALATLRKIMTALKEHQVDFRDIIFYFYS 180
QY 121 TFLQLERKGVSSGIMLTFWLVALVCALATLRKIMTALKEHQVDFRDIIFYFYS 180
DB 121 TFLQLERKGVSSGIMLTFWLVALVCALATLRKIMTALKEHQVDFRDIIFYFYS 180
QY 181 LLLIQLVLSCFSDRSLPSETIHDPNCPRESSASFLSRITFWITGLIVRGYRQPLEGSD 240
DB 181 LLLIQLVLSCFSDRSLPSETIHDPNCPRESSASFLSRITFWITGLIVRGYRQPLEGSD 240
QY 181 LLLIQLVLSCFSDRSLPSETIHDPNCPRESSASFLSRITFWITGLIVRGYRQPLEGSD 240
DB 181 LLLIQLVLSCFSDRSLPSETIHDPNCPRESSASFLSRITFWITGLIVRGYRQPLEGSD 240
QY 241 LMSLNKETSEOVVVLVYKMKKECAKTRKQPVKYVYS-SKDPAPKSSSKVDANEVEYA 299
DB 241 LMSLNKETSEOVVVLVYKMKKECAKTRKQPVKYVYS-SKDPAPKSSSKVDANEVEYA 299
QY 241 LMSLNKETSEOVVVLVYKMKKECAKTRKQPVKYVYS-SKDPAPKSSSKVDANEVEYA 299
DB 241 LMSLNKETSEOVVVLVYKMKKECAKTRKQPVKYVYS-SKDPAPKSSSKVDANEVEYA 299
QY 300 LVIKSPKEMNPSLSEKVLKTKGPFYLMSEFFKALHDLMMFSGPOLIKLLIKFVNDTYAP 359
DB 300 LVIKSPKEMNPSLSEKVLKTKGPFYLMSEFFKALHDLMMFSGPOLIKLLIKFVNDTYAP 359
QY 301 LIVSSHKDRDPSLKVLYKTKGPFYLMSEFFKALHDLMMFSGPOLIKLLIKFVNDTYAP 360
DB 301 LIVSSHKDRDPSLKVLYKTKGPFYLMSEFFKALHDLMMFSGPOLIKLLIKFVNDTYAP 360
QY 360 DMGGYFVTVLLFVTACLOTVLVHQYFHIQFVSGMBIKTAVIGAVYRKALVITNSARKSST 419
DB 360 DMGGYFVTVLLFVTACLOTVLVHQYFHIQFVSGMBIKTAVIGAVYRKALVITNSARKSST 419

DB 361 DMGGYFVTVLLFVTACLOTVLVHQYFHIQFVSGMBIKTAVIGAVYRKALVITNSARKSST 420
QY 420 VGEIVNMSVDAORFMDLATIYINMWSAPLOVILATLMLNIGPSYIAGAVVAVLVV 479
DB 421 VGEIVNMSVDAORFMDLATIYINMWSAPLOVILATLMLNIGPSYIAGAVVAVLVV 480
QY 480 NAWAMTKTYOVAHMSKDNRIKLANEINGIKVLKLYAMELAFKDKVLAIRQELKVL 539
DB 481 NAWAMTKTYOVAHMSKDNRIKLANEINGIKVLKLYAMELAFKDKVLAIRQELKVL 540
QY 540 KKSAYLSAVGTFMTWVCPPLVALCTFAVYVYTIENNLIDQOTAFVSLAFNIIARFPNII 599
DB 541 KKSAYLSAVGTFMTWVCPPLVALCTFAVYVYTIENNLIDQOTAFVSLAFNIIARFPNII 600
QY 600 PMVSSIVQSVSLKRIPLFSHELEPDSIERPVQDGGTNSITVRNATFTWASDDP 659
DB 601 PMVSSIVQSVSLKRIPLFSHELEPDSIERPVQDGGTNSITVRNATFTWASDDP 660
QY 660 TUNGITSEIGALVAVVGCGKSSLSALLAEMDKVEGHVAIKGSVAVVPOQAMION 719
DB 661 TUNGITSEIGALVAVVGCGKSSLSALLAEMDKVEGHVAIKGSVAVVPOQAMION 720
QY 720 DSIAPENILFGQLEEPYRYSVIOACALLPDLIELPSGDRTEIGRGVNLGGGQKQVSLA 779
DB 721 DSIAPENILFGQLEEPYRYSVIOACALLPDLIELPSGDRTEIGRGVNLGGGQKQVSLA 780
QY 780 RAVSNADLYLFDPLSLAVDAVHAKHIFENVYIGPKGMKNKTRILVTHMSYLPQVAVII 839
DB 781 RAVSNADLYLFDPLSLAVDAVHAKHIFENVYIGPKGMKNKTRILVTHMSYLPQVAVII 840
QY 840 VMSGKISEMGSYOELLARDGAPAEPLRTYASTEQEODABENGVTVSGGPKAKOMENG 899
DB 841 VMSGKISEMGSYOELLARDGAPAEPLRTYASTEQEODABENGVTVSGGPKAKOMENG 900
QY 900 MLVYDSAGKQORQLSSSSSYSGDISRHNSSTALQKAEKKEETWKLMEADKQOTGVK 959
DB 901 MLVYDSAGKQORQLSSSSSYSGDISRHNSSTALQKAEKKEETWKLMEADKQOTGVK 960
QY 960 LSVYVMDVWKAIGLIFISFISFLFPMCNHVSALASNYWLSMTD-DPIVNGTOEHTKVL 1018
DB 961 LSVYVMDVWKAIGLIFISFISFLFPMCNHVSALASNYWLSMTD-DPIVNGTOEHTKVL 1019
QY 1019 YGALGISQGIAPVGYMAVSIIGILASRCLHVDLHSILNSPMSFFERTPSGILVNRPSK 1078
DB 1020 YGALGISQGIAPVGYMAVSIIGILASRCLHVDLHSILNSPMSFFERTPSGILVNRPSK 1079
QY 1079 ELDTVDNMTPEVITKMPGSLFNVIAGCIVILLAPRIAIIIPRGLITYFFQRPVVASR 1138
DB 1080 ELDTVDNMTPEVITKMPGSLFNVIAGCIVILLAPRIAIIIPRGLITYFFQRPVVASR 1139
QY 1139 QLKRLSEVSRSPVYSHFNETLLGVSVIRAFEBQERFTHQSDLKVDENOKAVYPSIVANRW 1198
DB 1140 QLKRLSEVSRSPVYSHFNETLLGVSVIRAFEBQERFTHQSDLKVDENOKAVYPSIVANRW 1199
QY 1199 LAVRLCEVGNCLVFAALFAVISRHSLSAGLVGLSVSISQVTTYLMLVMSSEMETNI 1258
DB 1200 LAVRLCEVGNCLVFAALFAVISRHSLSAGLVGLSVSISQVTTYLMLVMSSEMETNI 1259
QY 1259 VAVRLKESYSTEKEAPWQIOETAPPSGWPQVGYVERPNTVCLTRREDLDFLRHINVTIN 1318
DB 1260 VAVRLKESYSTEKEAPWQIOETAPPSGWPQVGYVERPNTVCLTRREDLDFLRHINVTIN 1319
QY 1319 GGEKVIGYRTGAKSSLTGLPFINSGAGEIIIIIDGINIKIGLHDLRFKITTIIPODPV 1378
DB 1320 GGEKVIGYRTGAKSSLTGLPFINSGAGEIIIIIDGINIKIGLHDLRFKITTIIPODPV 1379
QY 1379 LFGSGLRNKNDLPFQOYSDDEEYVTSILELAHNDPVSAIPDKLDHCEACGEGNLSVGQRLV 1438
DB 1380 LFGSGLRNKNDLPFQOYSDDEEYVTSILELAHNDPVSAIPDKLDHCEACGEGNLSVGQRLV 1439
QY 1439 CLARALLKTKILVDEATAVVDLETDLDIOSITRTQEDCTVLTIAHRLNTINDYTRVI 1498
DB 1440 CLARALLKTKILVDEATAVVDLETDLDIOSITRTQEDCTVLTIAHRLNTINDYTRVI 1499

Qy 1499 VLDKGEIOEYGAAPSDLLQORGLFYSMADAGLV 1531
Db 1500 VLDKGEIRECAPSEBLLQORGFYSMAKADAGLV 1532

RESULT 13
Q8CG09 RAT PRELIMINARY; PRT; 1532 AA.
ID Q8CG09_RAT PRELIMINARY; PRT; 1532 AA.

AC -01-MAR-2003 (TREMBLrel. 23, Created)
DT -01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Multidrug resistance-associated protein 1.
GN Name=Abcc1; Synonyms=Mrp1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RA Yang Z., Li C.S.W., Shen D.D., Ho R.J.Y.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY170916; AAN86532.1; -; mRNA.
DR HSSP; P08716; LMT0.
DR Ensembl; ENSRNOG0000032748; Rattus norvegicus.
DR RGD; 3112; Abcc1.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR001395; Aldo/ket_red.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS00929; ABC_TMIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00063; ALDOXETO_REDUCTASE_3; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1532 AA; 171558 MW; C18F3554D85732A CRC64;

Query Match 70.0%; Score 6932.5; DB 2; Length 1532;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 1338; Conservative 100; Mismatches 92; Indels 3; Gaps 3;

Qy 1 MALGFCASDSDPLDMNMTNTSNPDKCFONTVLVWVPCFYLMACPFYFLYSRH 60
Db 1 MALRSPSSDSDPLDMNMTNTSNPDKCFONTVLVWVPCFYLMACPFYFLYSRH 60

Qy 61 DRGIIQMTPLNKTALGFLIIVCMADLFYSWERSGIFLAFVLSPTLLGITLLA 120
Db 61 DRGIIQMTPLNKTALGFLIIVCMADLFYSWERSGIFLAFVLSPTLLGITLLA 120

Qy 121 TELIQLRRKGVSSGIMLTFWLVAIVCALAIIKRSKIMTAKEDAOVDLPEDIFYYVFS 180
Db 121 TELIQLRRKGVSSGIMLTFWLVAIVCALAIIKRSKIMTAKEDAOVDLPEDIFYYVFS 180

Qy 121 TFLIQLRRKGVSSGIMLTFWLVAIVCALAIIKRSKIMTAKEDAOVDLPEDIFYYVFS 180
Db 121 TFLIQLRRKGVSSGIMLTFWLVAIVCALAIIKRSKIMTAKEDAOVDLPEDIFYYVFS 180

Qy 181 LLLIQLVLSGFSRSPPLFSETIHDNPNCPESSASFLSRITFWITGLIVRGYRPLSGSD 240
Db 181 LVIQLVLSGFSRSPPLFSETIHDNPNCPESSASFLSRITFWITGLIVRGYRPLSGSD 240

Qy 241 LMSLNKEDTSBOVVPVLVXNMKKECAKTRORPVXVVS -SKDPAQKSSKVDNEEVEA 299
Db 241 LMSLNKEDTSBOVVPVLVXNMKKECAKTRORPVXVVS -SKDPAQKSSKVDNEEVEA 299

Qy 300 LIYVSPKEMNPSLFKVLVYKTFPGYFLMSFFFKAIHDLMMFSGPOLKLLIKFVNDPKAP 359
Db 301 LIYVSHKDRDPSLFLKVLVYKTFPGYFLMSFLYKALHDLMMFAGEIIELEIINFNDREAP 360

Qy 360 DMQGYFTYVLLFTYACIQLTVLHQYFHIQVSGMRITAVIYAVYRRALVITNSARXST 419
Db 361 DMQGYFTYVLLFTYACIQLTVLHQYFHIQVSGMRITAVIYAVYRRALVITNSARXST 420

Qy 420 VGEIVNLSVDAQRFMDLATYIMINWAPQVLTALFLMLNIGPSVLAQAVWILNVPF 480
Db 421 VGEIVNLSVDAQRFMDLATYIMINWAPQVLTALFLMLNIGPSVLAQAVWILNVPF 480

Qy 480 NAWAMAKTKTYQVAHMKSKNRIKLMNEILNGIKVLKLYAMWELAFKDKVLAIROBELKVL 539
Db 481 NAWAMAKTKTYQVAHMKSKNRIKLMNEILNGIKVLKLYAMWELAFKDKVLAIROBELKVL 540

Qy 540 KKSAYLSAVGTFTWCTPPLVALCTPAVYVITDENNILDQAQFVSLALFNILRPLNLT 599
Db 541 KKSAYLSAVGTFTWCTPPLVALCTPAVYVITDENNILDQAQFVSLALFNILRPLNLT 600

Qy 600 PMYTSSTVQASVSLKRLRITLSHELEPDSIERPVQDGGTNSITRANITFWASDDP 659
Db 601 PMYTSSTVQASVSLKRLRITLSHELEPDSIERPVQDGGTNSITRANITFWASDDP 660

Qy 660 TLNGITSIPEGALVAVVGQVGGCKSLALALAEQDKVGAHAKSVAVYVPOQAVION 719
Db 661 TLNGITSIPEGALVAVVGQVGGCKSLALALAEQDKVGAHAKSVAVYVPOQAVION 720

Qy 720 DSIURENITLFCQLEEPYRSVIOACALLPLEILPESGDRTEIEGKVNLSCGQKORVSLA 779
Db 721 DSIURENITLFCQLEEPYRSVIOACALLPLEILPESGDRTEIEGKVNLSCGQKORVSLA 780

Qy 780 RAYYSNADITLFPDPLSAVDAAHVGKIFENVIGPKMKKKTITLVTHSMYSYLPQVDVIL 839
Db 781 RAYYSNADITLFPDPLSAVDAAHVGKIFENVIGPKMKKKTITLVTHSMYSYLPQVDVIL 840

Qy 840 VMSGKTSSEMSYQELIARGAFAEFLRTASTBOEDADENGVTGVSQKGAQKQWENG 899
Db 841 VMSGKTSSEMSYQELIARGAFAEFLRTASTBOEDADENGVTGVSQKGAQKQWENG 900

Qy 900 MLVTSAGKOLQORLSSSSSSSGDISRHNSSTAELOKAEAKKEETWLMADRAQOTQV 959
Db 901 MLVTSAGKOLQORLSSSSSSSGDISRHNSSTAELOKAEAKKEETWLMADRAQOTQV 959

Qy 960 LSYVMDYMKALIGLFIPLSLIFLMCNHVSALASNYMLSLMTDD-PIYNGTQEHKVLASV 1018
Db 960 LSYVMDYMKALIGLFIPLSLIFLMCNHVSALASNYMLSLMTDD-PIYNGTQEHKVLASV 1019

Qy 1019 YGALGTSOGIYAVFGYSAVSIIGIILASRCALVLLHSILRSPSFEERTSGNLYNRFK 1078
Db 1020 YGALGTSOGIYAVFGYSAVSIIGIILASRCALVLLHSILRSPSFEERTSGNLYNRFK 1079

Qy 1079 ELDTVDSMIPEVIMKMGSLFENVIGACIVILATPIAIIIPPLGLIYFFVQRFYVASSR 1138
Db 1080 ELDTVDSMIPEVIMKMGSLFENVIGACIVILATPIAIIIPPLGLIYFFVQRFYVASSR 1139

Qy 1139 QLRLESVSRSPIYVSHNEFTLLGVSVIRAFEEORPIHOSDLKYDENOKAYYSIVANRW 1198
Db 1140 QLRLESVSRSPIYVSHNEFTLLGVSVIRAFEEORPIHOSDLKYDENOKAYYSIVANRW 1199

Qy 1199 LAYRLCQVGCIVLFPALPAVISRHSLSAGLVGLSVSYSLQVTTYLNWILVRMSEMETNI 1258
Db 1200 LAYRLCQVGCIVLFPALPAVISRHSLSAGLVGLSVSYSLQVTTYLNWILVRMSEMETNI 1259

Qy 1259 VAVERLKEYSETERKEAPWQIOETAPPSQWQVGRVVEFRNYCLARYRBDLVLRHINVTIN 1318
Db 1260 VAVERLKEYSETERKEAPWQIOETAPPSQWQVGRVVEFRNYCLARYRBDLVLRHINVTIN 1319

QY 1319 GGEKVGIVGRTAGKSSITLGLFRINESAGEBIIIDGINIAKIGHADIRKITTIIIPDDPV 1378
DB 1320 GGEKVGIVGRTAGKSSITLGLFRINESAGEBIIIDGINIAKIGHADIRKITTIIIPDDPV 1379
QY 1379 LFGSLRWNLDPFGQYSDSEEWMTSLAHLKDPFSALEPDKLDHEGAGEMLSGQRPOLV 1438
DB 1380 LFGSLRWNLDPFGQYSDSEEWMTSLAHLKDPFSALEPDKLDHEGAGEMLSGQRPOLV 1439
QY 1439 CLARALLRKTKILVLEDEATAVDELTDLLIOSTRTQFEDCTVLTIARLNTIMDYTRVI 1498
DB 1440 CLARALLRKTKILVLEDEATAVDELTDLLIOSTRTQFEDCTVLTIARLNTIMDYTRVI 1499
QY 1499 VLDKGEIQEYCAPSDLLQQRGLFYSMAKADGLV 1531
DB 1500 VLDKGEIQEYCAPSDLLQQRGLFYSMAKADGLV 1532
RESULT 14
Q810G9 RAT PRELIMINARY; PRT; 1523 AA.
ID Q810G9 RAT PRELIMINARY; PRT; 1523 AA.
AC Q810G9;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE ATP-binding cassette protein C1 variant A.
GN Name=Abcc1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Spleen;
RA Takayanagi S., Ishikawa T.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBSJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY174892; AA044983.1; -; mRNA.
DR HSSP; P08716; IMTO.
DR RGD; 3112; Abcc1.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp1.
DR InterPro; IPR003439; ABC_transp1.
DR InterPro; IPR001395; Aldo/kec_red.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; Abc_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS00929; ABC_TM1F; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00063; ALDOKEETO_REDUCTASE_3; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1523 AA; 170505 MW; E40337051A1CB96 CRC64;
Query Match 69.6%; Score 6892; DB 2; Length 1523;
Best Local Similarity 87.0%; Pred. No. 0;
Matches 1333; Conservative 98; Mismatches 90; Indels 12; Gaps 4;
QY 1 MALRGFCAGDSDDPLMDNNVTNSNPDTKCFONTVYVWVPCFLYMACFPYFLYLSRH 60
DB 1 MALRSFCSSDSDPLMDNNVTNSNPDTKCFONTVYVWVPCFLYMACFPYFLYLSRH 60
QY 61 DRGYIQMTPLNKTALGFLMLIVCMADLFYSFWERSRGIFLAPVFLVSPFLIGITTLA 120

DB 61 DRGYIQMTPLNKTALGFLMLIVCMADLFYSFWERSRGIFLAPVFLVSPFLIGITTLA 120
QY 121 TFLIOLERRKGVSSGIMLTFMWALVACALALISKMTALKEDAQVDLDFRITFYVYS 180
DB 121 TFLIOLERRKGVSSGIMLTFMWALVACALALISKMTALKEDAQVDLDFRITFYVYS 180
QY 121 TFLIOLERRKGVSSGIMLTFMWALVACALALISKMTALKEDAQVDLDFRITFYVYS 180
DB 121 TFLIOLERRKGVSSGIMLTFMWALVACALALISKMTALKEDAQVDLDFRITFYVYS 180
QY 181 LLLIQLVLCSESDRSLFSETIHDNCPSESSASFLSRITFMWITGLIVRGYRPLEGSD 240
DB 181 LLLIQLVLCSESDRSLFSETIHDNCPSESSASFLSRITFMWITGLIVRGYRPLEGSD 240
QY 181 LLLIQLVLCSESDRSLFSETIHDNCPSESSASFLSRITFMWITGLIVRGYRPLEGSD 240
DB 181 LLLIQLVLCSESDRSLFSETIHDNCPSESSASFLSRITFMWITGLIVRGYRPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVYKWKKECACTRKQPVVYVS-SKDPKQPKSSVVDANEVEA 299
DB 241 LMSLNKEDTSEQVVPVYKWKKECACTRKQPVVYVS-SKDPKQPKSSVVDANEVEA 299
QY 241 LMSLNKEDTSEQVVPVYKWKKECACTRKQPVVYVS-SKDPKQPKSSVVDANEVEA 299
DB 241 LMSLNKEDTSEQVVPVYKWKKECACTRKQPVVYVS-SKDPKQPKSSVVDANEVEA 299
QY 300 LSVKSHKDRDPSLFKVLKTFGYPFLMSFFKAIHDLMPSPQILKLFVNDTKAP 359
DB 300 LSVKSHKDRDPSLFKVLKTFGYPFLMSFFKAIHDLMPSPQILKLFVNDTKAP 359
QY 301 LSVKSHKDRDPSLFKVLKTFGYPFLMSFFKAIHDLMPSPQILKLFVNDTKAP 360
DB 301 LSVKSHKDRDPSLFKVLKTFGYPFLMSFFKAIHDLMPSPQILKLFVNDTKAP 360
QY 360 DMQGYFTVLLFVTACIQLTVLHQYFHIQVSGNRKRTAVIGAVYRKALVITNSARKST 419
DB 360 DMQGYFTVLLFVTACIQLTVLHQYFHIQVSGNRKRTAVIGAVYRKALVITNSARKST 419
QY 361 DMQGYFTVLLFVTACIQLTVLHQYFHIQVSGNRKRTAVIGAVYRKALVITNSARKST 420
DB 361 DMQGYFTVLLFVTACIQLTVLHQYFHIQVSGNRKRTAVIGAVYRKALVITNSARKST 420
QY 420 VGEIIVNLSVDAQRFMDLATYINMINSAPLOVITLALYLMNIGPSVLAGVAVWLVMPV 479
DB 420 VGEIIVNLSVDAQRFMDLATYINMINSAPLOVITLALYLMNIGPSVLAGVAVWLVMPV 479
QY 421 VGEIIVNLSVDAQRFMDLATYINMINSAPLOVITLALYLMNIGPSVLAGVAVWLVMPV 480
DB 421 VGEIIVNLSVDAQRFMDLATYINMINSAPLOVITLALYLMNIGPSVLAGVAVWLVMPV 480
QY 480 NAWAMKTQYQVAMHMSKDNRIKLMEILNGIKYKLYAMELAFKQVLAIRQBELKVL 539
DB 480 NAWAMKTQYQVAMHMSKDNRIKLMEILNGIKYKLYAMELAFKQVLAIRQBELKVL 539
QY 481 NAWAMKTQYQVAMHMSKDNRIKLMEILNGIKYKLYAMELAFKQVLAIRQBELKVL 540
DB 481 NAWAMKTQYQVAMHMSKDNRIKLMEILNGIKYKLYAMELAFKQVLAIRQBELKVL 540
QY 540 KKSAYLSAVGTFTWCTPFLVALCTPAVYVITIDENIILDAQFVSLALENLRPLNITL 599
DB 540 KKSAYLSAVGTFTWCTPFLVALCTPAVYVITIDENIILDAQFVSLALENLRPLNITL 599
QY 541 KKSAYLSAVGTFTWCTPFLVALCTPAVYVITIDENIILDAQFVSLALENLRPLNITL 600
DB 541 KKSAYLSAVGTFTWCTPFLVALCTPAVYVITIDENIILDAQFVSLALENLRPLNITL 600
QY 600 PMVTSIYQASVSLKRLRIFLSHELEPDSIERPVPVDDGGTNSITTRNAFTTARSDPP 659
DB 600 PMVTSIYQASVSLKRLRIFLSHELEPDSIERPVPVDDGGTNSITTRNAFTTARSDPP 659
QY 601 PMVTSIYQASVSLKRLRIFLSHELEPDSIERPVPVDDGGTNSITTRNAFTTARSDPP 660
DB 601 PMVTSIYQASVSLKRLRIFLSHELEPDSIERPVPVDDGGTNSITTRNAFTTARSDPP 660
QY 660 TLNGITFIPGALVAVVGVGCGKSSLSLALMEMKVEGVAKSVAVPOQAMION 719
DB 660 TLNGITFIPGALVAVVGVGCGKSSLSLALMEMKVEGVAKSVAVPOQAMION 719
QY 661 TLNGITFIPGALVAVVGVGCGKSSLSLALMEMKVEGVAKSVAVPOQAMION 720
DB 661 TLNGITFIPGALVAVVGVGCGKSSLSLALMEMKVEGVAKSVAVPOQAMION 720
QY 720 DSLRENTLFGCOLPEPYSRYVQACALLPDLIELPSGDREIIGKGVNLSGGQKORVSLA 779
DB 720 DSLRENTLFGCOLPEPYSRYVQACALLPDLIELPSGDREIIGKGVNLSGGQKORVSLA 779
QY 721 DSLRENTLFGCOLPEPYSRYVQACALLPDLIELPSGDREIIGKGVNLSGGQKORVSLA 780
DB 721 DSLRENTLFGCOLPEPYSRYVQACALLPDLIELPSGDREIIGKGVNLSGGQKORVSLA 780
QY 780 RAVYSNADITYLPDPLSAVDAHVGKHIFENVYIGKMKLKKTRILLVTHSMYLPQVDYII 839
DB 780 RAVYSNADITYLPDPLSAVDAHVGKHIFENVYIGKMKLKKTRILLVTHSMYLPQVDYII 839
QY 781 RAVYSNADITYLPDPLSAVDAHVGKHIFENVYIGKMKLKKTRILLVTHSMYLPQVDYII 840
DB 781 RAVYSNADITYLPDPLSAVDAHVGKHIFENVYIGKMKLKKTRILLVTHSMYLPQVDYII 840
QY 840 VMSGKTSSEMSYQELLARDGAPAEPLRTYASTEOBDAEENGVYGVSGPKKAKOMENG 899
DB 840 VMSGKTSSEMSYQELLARDGAPAEPLRTYASTEOBDAEENGVYGVSGPKKAKOMENG 899
QY 841 VMSGKTSSEMSYQELLARDGAPAEPLRTYASTEOBDAEENGVYGVSGPKKAKOMENG 900
DB 841 VMSGKTSSEMSYQELLARDGAPAEPLRTYASTEOBDAEENGVYGVSGPKKAKOMENG 900
QY 900 MLVYDSAGKQOROLSSSSSYSGDISRHNSSTAELQKAEAKKETBWKLMEDAKQGTGVK 959
DB 900 MLVYDSAGKQOROLSSSSSYSGDISRHNSSTAELQKAEAKKETBWKLMEDAKQGTGVK 959
QY 901 ILVYDVGKPL-----HSVYTNQHSSTALEQSS-GVKEFTWKLMEADKXQGTGVK 950
DB 901 ILVYDVGKPL-----HSVYTNQHSSTALEQSS-GVKEFTWKLMEADKXQGTGVK 950
QY 960 LSVYWDYKAIQLFISPLSIFLPMONHVSALASNYWSLWTD-PIYNGTOEHTKRLSV 1018
DB 960 LSVYWDYKAIQLFISPLSIFLPMONHVSALASNYWSLWTD-PIYNGTOEHTKRLSV 1018
QY 951 LSVYWNWYKAIQLFISPLSIFLPMONHVSALASNYWSLWTD-PAVAVGTGJENENFRLSV 1010
DB 951 LSVYWNWYKAIQLFISPLSIFLPMONHVSALASNYWSLWTD-PAVAVGTGJENENFRLSV 1010
QY 1019 YGALGISGIVAPGVSNMAYVSGILASCRHVDLHSLTSLRSPMSFPEPTPSGNTVNRPSK 1078
DB 1019 YGALGISGIVAPGVSNMAYVSGILASCRHVDLHSLTSLRSPMSFPEPTPSGNTVNRPSK 1078
QY 1011 YGALGISGIVAPGVSNMAYVSGILASCRHVDLHSLTSLRSPMSFPEPTPSGNTVNRPSK 1070
DB 1011 YGALGISGIVAPGVSNMAYVSGILASCRHVDLHSLTSLRSPMSFPEPTPSGNTVNRPSK 1070
QY 1079 ELDTVDSMTPEVIKMFMSGLFNVIGACTIVILATPIAIIIPPLGLIYFFVQRFVASSR 1138
DB 1079 ELDTVDSMTPEVIKMFMSGLFNVIGACTIVILATPIAIIIPPLGLIYFFVQRFVASSR 1138
QY 1071 ELDTVDSMTPEVIKMFMSGLFNVIGACTIVILATPIAIIIPPLGLIYFFVQRFVASSR 1130
DB 1071 ELDTVDSMTPEVIKMFMSGLFNVIGACTIVILATPIAIIIPPLGLIYFFVQRFVASSR 1130
QY 1139 QLRLESVSRSPPVSHNETLLGVSVIRAFEEQRFHQSLKVDENOKAYYPSIIVANRW 1198
DB 1139 QLRLESVSRSPPVSHNETLLGVSVIRAFEEQRFHQSLKVDENOKAYYPSIIVANRW 1198

Dh 1131 QLRLESVSRSPVSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENOKAVYPSIVANRW 1190
Qy 1199 LAVRECEGNCIVFPAALFAVISRHSLSAGLVGSVSLQVTTVMNLVMSSEMENI 1258
Db 1191 LAVRECEGNCIVFPAALFAVISRHSLSAGLVGSVSLQVTTVMNLVMSSEMENI 1250
Qy 1259 VAVERLKEYSBTEKEAPMOIOETAPSSMPQGVAFEPNYCLARYEDLDFYLRHINTIN 1318
Db 1251 VAVRELKESBTEKEASMOIQETAPSPCTWPSGRFEFDYCLRYEDLDVLRKHINTIE 1310
Qy 1319 GGEKVGIVRTGAGKSLTTLGLFRINESAGEEIIIDGINAKIGHHLRPFKTTIIIPQDPV 1378
Db 1311 GGEKVGIVRTGAGKSLTTLGLFRINESAGEEIIIDGINAKIGHHLRPFKTTIIIPQDPV 1370
Qy 1379 LFGSLRNLPDPFGQYSDSEEWTSLEHLKDFVSALDDKIDHECAEGENLVSQGRQV 1438
Db 1371 LFGSLRNLPDPFGQYSDSEEWTSLEHLKDFVSALDDKIDHECAEGENLVSQGRQV 1430
Qy 1439 CLARALLRKTILVDEATAVDETDLTIQSTIRPEDCTVLTIARLNTIMDYTRVI 1498
Db 1431 CLARALLRKTILVDEATAVDETDLTIQSTIRPEDCTVLTIARLNTIMDYTRVI 1490
Qy 1499 VLDKGEIOEGAPSDLLQORGLFYSMAXDAGLV 1531
Db 1491 VLDKGEIRECGAPSELQORGVFYSMAXDAGLV 1523

RESULT 15
Q5F364_CHICK
ID Q5F364_CHICK PRELIMINARY; PRT; 1525 AA.
AC Q5F364;
DT 10-MAY-2005 (Tremblrel. 30, Created)
DT 10-MAY-2005 (Tremblrel. 30, Last sequence update)
DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=RCJMB04.32d20;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,
RA Plachy J., Kuter S., Blagodatzi A., Kostovska D., Kotler M.,
RA "Full-length cDNAs from chicken bursa lymphocytes to facilitate
RT gene function analysis";
RL Genome Biol. 6:R6-R6(2005).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AJ851786; CAH5420.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMIF_2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.

KW ATP-binding; Hypothetical protein; Nucleotide-binding; Repeat;
KW Transport. 1525 AA; 170972 MW; 3872486711B878FD CR664;
SQ SEQUENCE
Query Match 63.6%; Score 6301; DB 2; Length 1525;
Best Local Similarity 77.6%; Pred. No. 0;
Matches 1193; Conservative 176; Mismatches 148; Indels 20; Gaps 7;
Qy 1 MALRGFSGADSDLDLNMWNTMTNSNDFPKCFQNTLVWPCRYLACPFYFLYLSRH 60
Db 1 MGIESLSADASEPFWMNLTWHENPDPFOCFQNTLVWPCRYLWCFYFLYLSRH 60
Qy 61 DRGYIOMTPLNKTALGFLMIYCMADLFPSFERSRGIFLAVFVLSPLIGITLLA 120
Db 61 DRGYIOMTPLNKTALGFLMIYCMADLFPSFERSQNTFRAPFLISPLVGLITLLA 120
Qy 121 TFLIQLRRKGVSSGIMLTFWLVALYCALALRSKIMTALKEDAQVDLFRDITFYVFS 180
Db 121 TFLIQLRRKGVSSGIMLTFWLVALYCALALRSKIMTALKEDAQVDLFRDITFYVFS 180
Qy 121 TFLIQLRRKGVSSGIMLTFWLVALYCALALRSKIMTALKEDAQVDLFRDITFYVFS 180
Db 121 TFLIQLRRKGVSSGIMLTFWLVALYCALALRSKIMTALKEDAQVDLFRDITFYVFS 180
Qy 181 LLLIQLVLSGSDRSPLFSETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLSGD 240
Db 181 LLLIQLVLSGSDRSPLFSETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLSGD 240
Qy 181 LLLIQLVLSGSDRSPLFSETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLSGD 240
Db 181 LLLIQLVLSGSDRSPLFSETIHDNPNCPSSASFLSRITFWITGLIVRGYRQPLSGD 240
Qy 241 LMSLNKEDTSEOVVPLVKQWKKECAKTRKOPVKVYSSKDPAPKESKVDAN----EE 296
Db 241 LMSLNKEDTSEOVVPLVKQWKKECAKTRKOPVKVYSSKDPAPKESKVDAN----EE 296
Qy 241 LMSLNKEDTSEOVVPLVKQWKKECAKTRKOPVKVYSSKDPAPKESKVDAN----EE 296
Db 241 LMSLNKEDTSEOVVPLVKQWKKECAKTRKOPVKVYSSKDPAPKESKVDAN----EE 296
Qy 297 VEALIVKSPQENNPSEFLKVLKTFPGYFLMSPFPAKIHDLMSGPOILKILKIVNDT 356
Db 297 VEALIVKSPQENNPSEFLKVLKTFPGYFLMSPFPAKIHDLMSGPOILKILKIVNDT 356
Qy 296 ABAIIKRPQSSSASLSKVLKTFPGYFLMSPFPAKIHDLMTGHEILKLINPVNKK 355
Db 296 ABAIIKRPQSSSASLSKVLKTFPGYFLMSPFPAKIHDLMTGHEILKLINPVNKK 355
Qy 357 KAPWQGYFTVLLFTFACIQTLYLHOYFHCYSGKRIKAVIGAVYKRALVITNSARK 416
Db 357 KAPWQGYFTVLLFTFACIQTLYLHOYFHCYSGKRIKAVIGAVYKRALVITNSARK 416
Qy 356 SAPWQGYFTVLLFTFACIQTLYLHOYFHCYSGKRIKAVIGAVYKRALVITNSARK 415
Db 356 SAPWQGYFTVLLFTFACIQTLYLHOYFHCYSGKRIKAVIGAVYKRALVITNSARK 415
Qy 417 SSTVGEIVNMSVDAQFMDLATYINMWSAPLOVILATLYLMTNGPSVLAGAVWVM 476
Db 417 SSTVGEIVNMSVDAQFMDLATYINMWSAPLOVILATLYLMTNGPSVLAGAVWVM 476
Qy 416 TSTVGEIVNMSVDAQFMDLATYINMWSAPLOVILATLYLMTNGPSVLAGAVWML 475
Db 416 TSTVGEIVNMSVDAQFMDLATYINMWSAPLOVILATLYLMTNGPSVLAGAVWML 475
Qy 477 VPVNAVAMKTKTYQVAMKSKDNRIKLAMEILNGIKVLYLAMELAFKQVILAROBEL 536
Db 477 VPVNAVAMKTKTYQVAMKSKDNRIKLAMEILNGIKVLYLAMELAFKQVILAROBEL 536
Qy 476 VPVNAVAMKTKTYQVAMKSKDNRIKLAMEILNGIKVLYLAMELAFKQVILAROBEL 535
Db 476 VPVNAVAMKTKTYQVAMKSKDNRIKLAMEILNGIKVLYLAMELAFKQVILAROBEL 535
Qy 537 KVLKKSAYLSAVGTFTVCTPFLVALCTFAVYVTTIDENNILDAQTAFLAENILRPPL 596
Db 537 KVLKKSAYLSAVGTFTVCTPFLVALCTFAVYVTTIDENNILDAQTAFLAENILRPPL 596
Qy 536 KVLKKSAYLSAVGTFTVCTPFLVALCTFAVYVTTIDENNILDAQTAFLAENILRPPL 595
Db 536 KVLKKSAYLSAVGTFTVCTPFLVALCTFAVYVTTIDENNILDAQTAFLAENILRPPL 595
Qy 597 NILPMVTSIVQASVSLKRLRFLSHBELPDSIERPVDGGSNTSITVRNATFTWARS 656
Db 597 NILPMVTSIVQASVSLKRLRFLSHBELPDSIERPVDGGSNTSITVRNATFTWARS 656
Qy 596 NILPMVTSIVQASVSLKRLRFLSHBELPDSIERPVDGGSNTSITVRNATFTWARS 653
Db 596 NILPMVTSIVQASVSLKRLRFLSHBELPDSIERPVDGGSNTSITVRNATFTWARS 653
Qy 657 DPPTLNGITSIEGALVAVAGVCGSKSLLSALLAEMDKVEGHVAKGSVAVVPOQAW 716
Db 657 DPPTLNGITSIEGALVAVAGVCGSKSLLSALLAEMDKVEGHVAKGSVAVVPOQAW 716
Qy 654 DPPTLNGITSIEGALVAVAGVCGSKSLLSALLAEMDKVEGHVAKGSVAVVPOQAW 713
Db 654 DPPTLNGITSIEGALVAVAGVCGSKSLLSALLAEMDKVEGHVAKGSVAVVPOQAW 713
Qy 717 IONDSLENTLIFGCGLEBPYRSVIOACALLPDLLEIIPSGDRTRIGEGYNLSGGQORV 776
Db 717 IONDSLENTLIFGCGLEBPYRSVIOACALLPDLLEIIPSGDRTRIGEGYNLSGGQORV 776
Qy 714 IONATLEDNIIIFGEMNESRYKRVIEACALLPDLLEIIPMDRTRIGEGYNLSGGQORV 773
Db 714 IONATLEDNIIIFGEMNESRYKRVIEACALLPDLLEIIPMDRTRIGEGYNLSGGQORV 773
Qy 777 SLRAVYSNADYLPDPLSAVDAHVKCHLFEENVIGKMKLKNKTRILVTHMSYLPQVD 836
Db 777 SLRAVYSNADYLPDPLSAVDAHVKCHLFEENVIGKMKLKNKTRILVTHMSYLPQVD 836
Qy 774 SLRAVYSNADYLPDPLSAVDAHVKCHLFEENVIGKMKLKNKTRILVTHMSYLPQVD 833
Db 774 SLRAVYSNADYLPDPLSAVDAHVKCHLFEENVIGKMKLKNKTRILVTHMSYLPQVD 833
Qy 837 VTIIVSGKISSEMSYOELIARDAFAEFRTYASTEOQDAEENGVTGVSQP-GKEAKQ 895
Db 837 VTIIVSGKISSEMSYOELIARDAFAEFRTYASTEOQDAEENGVTGVSQP-GKEAKQ 895
Qy 834 TIIIVTGEISEMSYQELIKQGAFAEFRTYANMNSNESSD-----ASSSGKSGKP 888
Db 834 TIIIVTGEISEMSYQELIKQGAFAEFRTYANMNSNESSD-----ASSSGKSGKP 888
Qy 896 MENGMVLTDAQOLQOLSSSSSYGSDI--SEHNSTALQRAEAKKETVTLMDKXA 953
Db 896 MENGMVLTDAQOLQOLSSSSSYGSDI--SEHNSTALQRAEAKKETVTLMDKXA 953
Qy 889 VENGVLVNDATGLKMLHQLSNSSYSSTRETKSQHQSSTAELQKPLAEK-NSMKLTEDTA 947
Db 889 VENGVLVNDATGLKMLHQLSNSSYSSTRETKSQHQSSTAELQKPLAEK-NSMKLTEDTA 947
Qy 954 QTGVQKLSVYVMDVKAIGLIFISFLFPCNHYVALASVYMLSLMTDPIVNGTQDHTK 1013
Db 954 QTGVQKLSVYVMDVKAIGLIFISFLFPCNHYVALASVYMLSLMTDPIVNGTQDHTK 1013

```

Db      948  KTGKAVKATVWEYMKAIGLYISFLSVFLMCMCHIASLASNYWLSLMTDDPVVNGTQOYTN 1007
QY      1014  VRLSVYGAIGISOGIAVGYSMVAISIGGILASRCIHLVDLHSILRSPMSFEERTPSGNLV 1073
Db      1008  VRLGVYGAIGISOGIAVGYSMVAISIGGILASRCIHLVDLHSILRSPMSFEERTPSGNLV 1067
QY      1074  NRSFKELDTVDSMIPEVIKMFMSLENNVIGACIVILATPIAIIIPPLGIYFPVQRFY 1133
Db      1068  SRSFKELDTVDSMIPEVIKMFMSLENNVIGACIVILATPIAIIIPPLGIYFPVQRFY 1127
QY      1134  VASSRQLKRLSVSRSPYSHNETLLGVSVIRAFEBQERFIHQSDLVKVDENOKAYYPSI 1193
Db      1128  VATSRQLKRLSVSRSPYSHNETLLGVSVIRAFEBQERFIHQSDLVKVDENOKAYYPSI 1187
QY      1194  VANRWLAVERLCEVNCIVLPALPAVIRSHSISAGLVGYSYSLOVTTYLNMLVRMSSSE 1253
Db      1188  VANRWLAVERLCEVNCIVLPALPAVIRSHSISAGLVGYSYSLOVTTYLNMLVRMSSSE 1247
QY      1254  METNIVAVEERLKEYSETEKEAPWQIQETAPPSWPQVGRVEFRNYCLRYREDDLVLRHI 1313
Db      1248  LETNIVAVEERLKEYSETEKEAPWQIQETAPPSWPQVGRVEFRNYCLRYREDDLVLRHI 1307
QY      1314  NVTINGGEKVGIVGRTGAGKSSLTGLFRINSAEGEIIIDGINIAKIGLHDLRPFKTI 1373
Db      1308  NVTINGGEKVGIVGRTGAGKSSLTGLFRINSAEGEIIIDGINIAKIGLHDLRPFKTI 1367
QY      1374  FODPVLFGSGLRMNLDPPSOYSEDEWMTSLLELAHLKDFVSALPDKLDHCAEGENLSVG 1433
Db      1368  FODPVLFGSGLRMNLDPPSOYSEDEWMTSLLELAHLKDFVSALPDKLDHCAEGENLSVG 1427
QY      1434  QROLVCLARALLRKTIIIVLDEATAVLDLETDLLIQSTIRIQFEDCTVLTIAHRLNTIMD 1493
Db      1428  QROLVCLARALLRKTIIIVLDEATAVLDLETDLLIQSTIRIQFEDCTVLTIAHRLNTIMD 1487
QY      1494  YTRVIVLDKGEIQEYGAPSDLIQORGLFYSMKXDAGL 1530
Db      1488  YTRVIVLDRGEVCECDSPDNLQAKGLFYSMKXDSGL 1524

```

Search completed: December 15, 2005, 15:33:24
 Job time : 257.254 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 15:05:25 ; Search time 38.351 Seconds
(without alignments)
4076.547 Million cell updates/sec

Title: US-10-665-283-4

Perfect score: 9734
Sequence: 1 MALRGGFCASGSDPLMDMNV.....NTIKVPTPLCTARQLDSDRS 1691

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/1/1aa/5_COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/6_COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/H_COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/R_COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7860	80.7	1531	1 US-08-463-092B-4	Sequence 4, Appli
2	7860	80.7	1531	1 US-08-462-109A-4	Sequence 4, Appli
3	7860	80.7	1531	1 US-08-460-907B-4	Sequence 4, Appli
4	7860	80.7	1531	2 US-08-463-179A-4	Sequence 4, Appli
5	7860	80.7	1531	2 US-08-461-384B-4	Sequence 4, Appli
6	7849	80.6	1531	2 US-09-647-140B-19	Sequence 19, Appli
7	7849	80.6	1531	1 US-08-141-893-2	Sequence 2, Appli
8	7849	80.6	1531	1 US-08-463-092B-2	Sequence 2, Appli
9	7849	80.6	1531	1 US-08-462-109A-2	Sequence 2, Appli
10	7849	80.6	1531	1 US-08-460-907B-2	Sequence 2, Appli
11	7849	80.6	1531	2 US-08-463-179A-2	Sequence 2, Appli
12	7849	80.6	1531	2 US-08-461-384B-2	Sequence 2, Appli
13	7849	80.6	1531	2 US-08-407-207A-2	Sequence 2, Appli
14	7002.5	71.9	1528	1 US-08-463-092B-6	Sequence 6, Appli
15	7002.5	71.9	1528	1 US-08-462-109A-6	Sequence 6, Appli
16	7002.5	71.9	1528	1 US-08-460-907B-6	Sequence 6, Appli
17	7002.5	71.9	1528	2 US-08-463-179A-6	Sequence 6, Appli
18	7002.5	71.9	1528	2 US-08-461-384B-6	Sequence 6, Appli
19	4487.5	46.1	1527	2 US-09-647-140B-6	Sequence 6, Appli
20	4487.5	46.1	1527	2 US-09-647-140B-33	Sequence 33, Appli
21	3391.5	34.8	1503	2 US-09-647-140B-8	Sequence 8, Appli
22	3385.5	34.8	1503	2 US-09-792-616-3	Sequence 3, Appli
23	3218.5	33.1	1498	2 US-09-792-616-9	Sequence 3, Appli
24	2374.5	24.4	1621	2 US-08-972-927-3	Sequence 3, Appli
25	2349.5	24.1	1622	2 US-08-972-927-6	Sequence 6, Appli
26	2266.5	23.3	1325	2 US-09-647-140B-2	Sequence 2, Appli
27	2212.5	22.7	1464	2 US-10-012-896-1008	Sequence 1008, Ap

ALIGNMENTS

28	2195.5	22.6	1261	2 US-09-439-313-538	Sequence 538, App
29	2195.5	22.6	1261	2 US-09-636-215-538	Sequence 538, App
30	2195.5	22.6	1261	2 US-09-685-166A-538	Sequence 538, App
31	2195.5	22.6	1261	2 US-09-679-426-538	Sequence 538, App
32	2195.5	22.6	1261	2 US-09-759-143-538	Sequence 538, App
33	2195.5	22.6	1261	2 US-09-651-236-538	Sequence 538, App
34	2195.5	22.6	1261	2 US-09-657-279-538	Sequence 538, App
35	2195.5	22.6	1261	2 US-10-012-896-538	Sequence 538, App
36	2195.5	22.6	1261	2 US-10-012-896-1009	Sequence 1009, App
37	2118	21.8	1581	2 US-08-726-520-3	Sequence 3, Appli
38	2118	21.8	1581	2 US-09-208-716-3	Sequence 3, Appli
39	2115	21.7	1228	2 US-09-439-313-537	Sequence 537, App
40	2115	21.7	1228	2 US-09-636-215-537	Sequence 537, App
41	2115	21.7	1228	2 US-09-685-166A-537	Sequence 537, App
42	2115	21.7	1228	2 US-09-679-426-537	Sequence 537, App
43	2115	21.7	1228	2 US-09-759-143-537	Sequence 537, App
44	2115	21.7	1228	2 US-09-651-236-537	Sequence 537, App
45	2115	21.7	1228	2 US-09-657-279-537	Sequence 537, App

RESULT 1
US-08-463-092B-4
Sequence 4, Application US/08463092B
Patent No. 5766880
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
MULTIDRUG RESISTANCE PROTEINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: PARTEO RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,092B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: Q1546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1531 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-463-0928-4

Query Match 80.7%; Score 7860; DB 1; Length 1531;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MALRFGSADGSDPLMDMNTWNTSNPDTFCQNTVLVWYPCFTLWACFPFFLYLSRH 60
DB 1 MALRFGSADGSDPLMDMNTWNTSNPDTFCQNTVLVWYPCFTLWACFPFFLYLSRH 60
QY 61 DRGIOMTPLNKKTALGELLMIVCMADLFYSFMRSRSGIFLAPFLVSPILLGTTLLA 120
DB 61 DRGIOMTPLNKKTALGELLMIVCMADLFYSFMRSRSGIFLAPFLVSPILLGTTLLA 120
QY 121 TFLQLERRKGVSSGIMLTFMLVALVICALILRSKIMTALKEDAQVDLFRDITFFVYFS 180
DB 121 TFLQLERRKGVSSGIMLTFMLVALVICALILRSKIMTALKEDAQVDLFRDITFFVYFS 180
QY 121 TFLQLERRKGVSSGIMLTFMLVALVICALILRSKIMTALKEDAQVDLFRDITFFVYFS 180
DB 121 TFLQLERRKGVSSGIMLTFMLVALVICALILRSKIMTALKEDAQVDLFRDITFFVYFS 180
QY 181 LLLIQLVLSGSDSPLEFSETIHDNPPCESSASFLSITTTWITGLIVRGYRQPLEGSD 240
DB 181 LLLIQLVLSGSDSPLEFSETIHDNPPCESSASFLSITTTWITGLIVRGYRQPLEGSD 240
QY 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKQPVKVYSSKDPAPCKESSKYDANEVVAL 300
DB 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKQPVKVYSSKDPAPCKESSKYDANEVVAL 300
QY 301 IVKSPQKEMNSPLFVLVKTGFPYFLMSFFPKAIDHLMFSGPOLIKLIRFVNDTKAPD 360
DB 301 IVKSPQKEMNSPLFVLVKTGFPYFLMSFFPKAIDHLMFSGPOLIKLIRFVNDTKAPD 360
QY 361 WQGFYTYLLVVTACLOTVLVHQFHICFVSGMRIKTAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYTYLLVVTACLOTVLVHQFHICFVSGMRIKTAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIVNMSVDARFMDLATYINMTSAPLOVTLAYLMLNIGSPVLAVGVAVMTVMVNVN 480
DB 421 GEIVNMSVDARFMDLATYINMTSAPLOVTLAYLMLNIGSPVLAVGVAVMTVMVNVN 480
QY 481 AVMAKTKTYOVAAHKSKDNRIKLMNEILNGIKVLKYAMELAFKQVLAIRQEBLKYLK 540
DB 481 AVMAKTKTYOVAAHKSKDNRIKLMNEILNGIKVLKYAMELAFKQVLAIRQEBLKYLK 540
QY 541 KSAVLSAVGTFTWCTPPLVALCTFAVYTTDENNILDQAQFVSLALFNILRPLNLTLP 600
DB 541 KSAVLSAVGTFTWCTPPLVALCTFAVYTTDENNILDQAQFVSLALFNILRPLNLTLP 600
QY 601 MVISSIVASVSIKRLRIFLSHEBELPDSIERRPVKDGGTMSITVRNATFTWASDPT 660
DB 601 MVISSIVASVSIKRLRIFLSHEBELPDSIERRPVKDGGTMSITVRNATFTWASDPT 660
QY 661 INGTFTSIPEGALVAVWGQVCGKSSLSALLAEMDKYEGHVAIKGSVAYVPOQAMIOND 720
DB 661 INGTFTSIPEGALVAVWGQVCGKSSLSALLAEMDKYEGHVAIKGSVAYVPOQAMIOND 720
QY 721 SLRENIIFGCQLEEBYRSVIOACALLPDLLEILPSGDRTEIGEKVNIISGQKQFVSLAR 780
DB 721 SLRENIIFGCQLEEBYRSVIOACALLPDLLEILPSGDRTEIGEKVNIISGQKQFVSLAR 780
QY 781 AVYSNADTYLPDDPLSAVDANHGKHIFENVIGPKMKLKKRILVTHSSSYLPQVDVIV 840
DB 781 AVYSNADTYLPDDPLSAVDANHGKHIFENVIGPKMKLKKRILVTHSSSYLPQVDVIV 840
QY 841 MSGKISMGSYOELLARDGAFABFLRTASTEOBDAEENGVTGSGGKEAKOMENGM 900
DB 841 MSGKISMGSYOELLARDGAFABFLRTASTEOBDAEENGVTGSGGKEAKOMENGM 900
QY 901 LVTSAGKOLOROLSSSSSYSGDISRHNSLAELOKAEAKKETWKLMEADXAQTGOYVL 960
DB 901 LVTSAGKOLOROLSSSSSYSGDISRHNSLAELOKAEAKKETWKLMEADXAQTGOYVL 960

```

```

DB 901 LVTSAGKOLOROLSSSSSYSGDISRHNSLAELOKAEAKKETWKLMEADXAQTGOYVL 960
QY 961 SVYDYMKAIGLFIPIPSLIFLMCNHVSALASNTWLSLMTDDPIVNGTOEHTKRLSYVG 1020
DB 961 SVYDYMKAIGLFIPIPSLIFLMCNHVSALASNTWLSLMTDDPIVNGTOEHTKRLSYVG 1020
QY 1021 ALGISGIAVFGYSMAVSIIGIILASRCLAYDLHLSILRSPMSFPERTPSGNLVNRFSKEL 1080
DB 1021 ALGISGIAVFGYSMAVSIIGIILASRCLAYDLHLSILRSPMSFPERTPSGNLVNRFSKEL 1080
QY 1081 DTUOSMTPEVYKMGMSLFENVIGACIVILATPIAIIIPPLGLIFFVQRFVYASSROL 1140
DB 1081 DTUOSMTPEVYKMGMSLFENVIGACIVILATPIAIIIPPLGLIFFVQRFVYASSROL 1140
QY 1141 KRLSVRSRPVYSHFNETLIGVSYIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLSVRSRPVYSHFNETLIGVSYIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLCEVNCIYLPALFAVISRHSLSAGLVGLSVSYLQVTTYINMTVRMSSEMETNIVA 1260
DB 1201 VRLCEVNCIYLPALFAVISRHSLSAGLVGLSVSYLQVTTYINMTVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMOIQETAPPSWPQVGRVFRNYCLRYREDLPVLRIHIVTTNGG 1320
DB 1261 VERLKEYSETEKEAPMOIQETAPPSWPQVGRVFRNYCLRYREDLPVLRIHIVTTNGG 1320
QY 1321 EKVGIVRTGAGKSLTGLFRINSAAGEIILIDGINIAKIGLHDLRFKTIITIPQDPYLF 1380
DB 1321 EKVGIVRTGAGKSLTGLFRINSAAGEIILIDGINIAKIGLHDLRFKTIITIPQDPYLF 1380
QY 1381 SGLRMLNDPFSQYSDDEWMTSLBLAHKDFVSAIPDKLDHECABGEGNLSVGRQOLVCL 1440
DB 1381 SGLRMLNDPFSQYSDDEWMTSLBLAHKDFVSAIPDKLDHECABGEGNLSVGRQOLVCL 1440
QY 1441 ARALLRKTILVDEATRAVDLETDLIQSTIRQFQEDCTVLTIAHRLNTIMDYTRIVL 1500
DB 1441 ARALLRKTILVDEATRAVDLETDLIQSTIRQFQEDCTVLTIAHRLNTIMDYTRIVL 1500
QY 1501 DKGEIOEYGAPSDILQORGLFYSMAXDAGLV 1531
DB 1501 DKGEIOEYGAPSDILQORGLFYSMAXDAGLV 1531

RESULT 2
US-08-462-109A-4
; Sequence 4, Application US/08462109A
; Patent No. 5882875
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,109A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340

```

FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: POI-002CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-4

Query Match 80.7%; Score 7860; DB 1; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCADGSDPLMDMNTMTSNPDFTKCFONTVLMVPCFYLMACFPFYFLYSRH 60
DB 1 MALRGFCADGSDPLMDMNTMTSNPDFTKCFONTVLMVPCFYLMACFPFYFLYSRH 60
QY 61 DRGYIQTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLAPVPLVSPLLGITTLLA 120
DB 61 DRGYIQTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLAPVPLVSPLLGITTLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFMLVALCALALISKIMTAKEDAOVDLPFDITFYFFS 180
DB 121 TFLIQLERRKGVSSGIMLTFMLVALCALALISKIMTAKEDAOVDLPFDITFYFFS 180
QY 121 TFLIQLERRKGVSSGIMLTFMLVALCALALISKIMTAKEDAOVDLPFDITFYFFS 180
DB 121 TFLIQLERRKGVSSGIMLTFMLVALCALALISKIMTAKEDAOVDLPFDITFYFFS 180
QY 181 LLLIQLVLSGSDSPFSETIHDNPNCPRESSASFLSRITFMWITGLIVRGYRQLGSD 240
DB 181 LLLIQLVLSGSDSPFSETIHDNPNCPRESSASFLSRITFMWITGLIVRGYRQLGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVVYSSKDPAPKSSKVDANEYEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVVYSSKDPAPKSSKVDANEYEAL 300
QY 301 IVSPQKEMNPSPLEKVTYKTFGPFYFMSFFPKAIDHLMFSGPOILKLIKFNVDTRAPD 360
DB 301 IVSPQKEMNPSPLEKVTYKTFGPFYFMSFFPKAIDHLMFSGPOILKLIKFNVDTRAPD 360
QY 361 WQGYFYVLLFEVTAQLOTLVLMHOYFHIQFVSGMRKTAIVGAVYRKALVITNSARKSSTV 420
DB 361 WQGYFYVLLFEVTAQLOTLVLMHOYFHIQFVSGMRKTAIVGAVYRKALVITNSARKSSTV 420
QY 421 GEIYNLMSVDAQRPMDLATYINMIMSAFLQVITLALVLLMNLGSPVLAVVWLAMPVN 480
DB 421 GEIYNLMSVDAQRPMDLATYINMIMSAFLQVITLALVLLMNLGSPVLAVVWLAMPVN 480
QY 481 AVMAMKTKTYQVAMKSKDNRIKLMNEILNGIKTLKAYAMELAFKDYLAIRQELVYLK 540
DB 481 AVMAMKTKTYQVAMKSKDNRIKLMNEILNGIKTLKAYAMELAFKDYLAIRQELVYLK 540
QY 541 KSAVLASVGTFTWCTPELVALCTEAYVYITDENNIIDAQAFVSLAFNLIRPLNLP 600
DB 541 KSAVLASVGTFTWCTPELVALCTEAYVYITDENNIIDAQAFVSLAFNLIRPLNLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHELEPDSIERRPVKDGGGINSITVNNATFTWARSDDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHELEPDSIERRPVKDGGGINSITVNNATFTWARSDDPT 660
QY 661 LNTGTFSPREGALVAVVGVCGGSSLSLALLAMDVYEGVAAKGSVAVYPOQAWIOND 720
DB 661 LNTGTFSPREGALVAVVGVCGGSSLSLALLAMDVYEGVAAKGSVAVYPOQAWIOND 720
QY 721 SLRENILFGCQLEBPYRSVIAOCALLPDEILPSGRTEIGEGVNLSSGQGRVSLAR 780

DB 721 SLRENILFGCQLEBPYRSVIAOCALLPDEILPSGRTEIGEGVNLSSGQGRVSLAR 780
QY 781 AAVSNADITYLPDPLSLAVDAVGHAFENYIGPFGMKAKTRILVYTHSMSTLPDQVYIIV 840
DB 781 AAVSNADITYLPDPLSLAVDAVGHAFENYIGPFGMKAKTRILVYTHSMSTLPDQVYIIV 840
QY 841 MSGGKISEMSYOELARDGAPAFELRTYASTBOEODAEENGVTGVSQPGKAAQOMENG 900
DB 841 MSGGKISEMSYOELARDGAPAFELRTYASTBOEODAEENGVTGVSQPGKAAQOMENG 900
QY 901 LVYDSAKQOLQORLSSSSSYSGDISRHNSTABLOKAEKETWKMMEADKAOTGVYL 960
DB 901 LVYDSAKQOLQORLSSSSSYSGDISRHNSTABLOKAEKETWKMMEADKAOTGVYL 960
QY 961 SVYWDYKAIQGLFISFLSIFLFCMNHVSALASNYWLSMTDDPIVNGTOERTKRLSVYG 1020
DB 961 SVYWDYKAIQGLFISFLSIFLFCMNHVSALASNYWLSMTDDPIVNGTOERTKRLSVYG 1020
QY 1021 ALGISOGIAVFGYMAVSIQILASRCLHVDLHSLIRSPMSFEKTPSGNLVNRFSKEL 1080
DB 1021 ALGISOGIAVFGYMAVSIQILASRCLHVDLHSLIRSPMSFEKTPSGNLVNRFSKEL 1080
QY 1081 DTVDMSIPEVIKMFMSGLFNVIGACIYILATPIAIIIPPLGIFYFVORFYAASSROL 1140
DB 1081 DTVDMSIPEVIKMFMSGLFNVIGACIYILATPIAIIIPPLGIFYFVORFYAASSROL 1140
QY 1141 KRLSVSRSPVSHFNELLGVSIVYRAFEQERFIHQSDLKVDENQAAVYPSIVANRWLA 1200
DB 1141 KRLSVSRSPVSHFNELLGVSIVYRAFEQERFIHQSDLKVDENQAAVYPSIVANRWLA 1200
QY 1201 VRLCVCNCIVLPALFAVISRSLSGVLGVSYSLOVTTYLNMVLVRRMSSEMETNIVA 1260
DB 1201 VRLCVCNCIVLPALFAVISRSLSGVLGVSYSLOVTTYLNMVLVRRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVFEFRYCYARREDLPVLRHINVTNGG 1320
DB 1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVFEFRYCYARREDLPVLRHINVTNGG 1320
QY 1321 EKVGIAGRTAGKSSLLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPLF 1380
DB 1321 EKVGIAGRTAGKSSLLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPLF 1380
QY 1381 SSGSRMNLDPFSQYSDEEFTSLBLAKPVSALPKLDHECAEGEENISVGOEOLVCL 1440
DB 1381 SSGSRMNLDPFSQYSDEEFTSLBLAKPVSALPKLDHECAEGEENISVGOEOLVCL 1440
QY 1441 ABALLRKTILVDEATAVADLETDLIQSTIRTOFEDCTVLTAHRLNTIMDYTRYIYL 1500
DB 1441 ABALLRKTILVDEATAVADLETDLIQSTIRTOFEDCTVLTAHRLNTIMDYTRYIYL 1500
QY 1501 DKGEIOEYAPSDLLQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYAPSDLLQORGLFYSMAKDAGLV 1531

RESULT 3
US-08-460-907B-4
Sequence 4, Application US/08460907B
Patent No. 5891724

GENERAL INFORMATION:
APPLICANT: Deeley, Roger G.
APPLICANT: Cole, Susan P.C.
TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG
RESISTANCE ON A CELL
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: PARTNO RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6

```

? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: ASCII text
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/460,907B
? FILING DATE: 05-JUN-1995
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/966,923
? FILING DATE: 27-OCT-1992
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/029,340
? FILING DATE: 8-MAR-1993
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/141,893
? FILING DATE: 26-OCT-1993
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/407,207
? FILING DATE: 20-MAR-1995
? CLASSIFICATION: 424
? ATTORNEY/AGENT INFORMATION:
? NAME: Steeg, Carol Miernicki
? REGISTRATION NUMBER: 39,539
? REFERENCE/DOCKET NUMBER: 01551
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (613) 545-2342
? TELEFAX: (613) 545-6853
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1531 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-460-907B-4

Query Match      80.7%; Score 7860; DB 1; Length 1531;
Beet local similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MALRGFCSDADSDPLMDNNTWNTSNPDFTKCFONTVLWVPCFYLMACFPEFYELYSRH 60
DB      1 MALRGFCSDADSDPLMDNNTWNTSNPDFTKCFONTVLWVPCFYLMACFPEFYELYSRH 60
QY      61 DRGYIQMTPLNKTKTALGFLIMVCMADLFYSFWEERSGIFLAPVFLVSPILLGITLLA 120
DB      61 DRGYIQMTPLNKTKTALGFLIMVCMADLFYSFWEERSGIFLAPVFLVSPILLGITLLA 120
QY      121 TFLIQLEBKRGVQSSGIMLTFWLVALVCAALILRSKITALKEDQVDFRDITFYVFS 180
DB      121 TFLIQLEBKRGVQSSGIMLTFWLVALVCAALILRSKITALKEDQVDFRDITFYVFS 180
QY      121 TFLIQLEBKRGVQSSGIMLTFWLVALVCAALILRSKITALKEDQVDFRDITFYVFS 180
DB      121 TFLIQLEBKRGVQSSGIMLTFWLVALVCAALILRSKITALKEDQVDFRDITFYVFS 180
QY      181 LLLIQVLVSCSDSPLESETIHDNPPCESASFLSRITFWMITGLIVRGYROPLEBSD 240
DB      181 LLLIQVLVSCSDSPLESETIHDNPPCESASFLSRITFWMITGLIVRGYROPLEBSD 240
QY      181 LLLIQVLVSCSDSPLESETIHDNPPCESASFLSRITFWMITGLIVRGYROPLEBSD 240
DB      181 LLLIQVLVSCSDSPLESETIHDNPPCESASFLSRITFWMITGLIVRGYROPLEBSD 240
QY      241 LMSLNKEDTSQVAVPVLVKMKKECAKTRKQPVKVYVSSKDPAPQKSSKYDANEVEAL 300
DB      241 LMSLNKEDTSQVAVPVLVKMKKECAKTRKQPVKVYVSSKDPAPQKSSKYDANEVEAL 300
QY      241 LMSLNKEDTSQVAVPVLVKMKKECAKTRKQPVKVYVSSKDPAPQKSSKYDANEVEAL 300
DB      241 LMSLNKEDTSQVAVPVLVKMKKECAKTRKQPVKVYVSSKDPAPQKSSKYDANEVEAL 300
QY      301 IVKSPQKEMNSPLFVLYKTGPFYLMSPFFKAIHDLMMFSGPQILKLIFVNDTKAPD 360
DB      301 IVKSPQKEMNSPLFVLYKTGPFYLMSPFFKAIHDLMMFSGPQILKLIFVNDTKAPD 360
QY      301 IVKSPQKEMNSPLFVLYKTGPFYLMSPFFKAIHDLMMFSGPQILKLIFVNDTKAPD 360
DB      301 IVKSPQKEMNSPLFVLYKTGPFYLMSPFFKAIHDLMMFSGPQILKLIFVNDTKAPD 360
QY      361 MGGFYTYLLEFVYACLOTLVLAHQVPHICFVSGMRIKTAVIGAVYKALVITNSAKKSTV 420
DB      361 MGGFYTYLLEFVYACLOTLVLAHQVPHICFVSGMRIKTAVIGAVYKALVITNSAKKSTV 420
QY      361 MGGFYTYLLEFVYACLOTLVLAHQVPHICFVSGMRIKTAVIGAVYKALVITNSAKKSTV 420
DB      361 MGGFYTYLLEFVYACLOTLVLAHQVPHICFVSGMRIKTAVIGAVYKALVITNSAKKSTV 420
QY      421 GEIVNLSVDAQRFMDLATYINIMWSAPLOVILALYLLMLNLGSPVLGVAVMVLMEVFN 480
```

```

DB      421 GEIVNLSVDAQRFMDLATYINIMWSAPLOVILALYLLMLNLGSPVLGVAVMVLMEVFN 480
QY      481 AVNMAKTKTYQVAHMKSKDRIRKLMNEILNGIKVLKLYAMELAFKDKVLAIRQBELKYLK 540
DB      481 AVNMAKTKTYQVAHMKSKDRIRKLMNEILNGIKVLKLYAMELAFKDKVLAIRQBELKYLK 540
QY      481 AVNMAKTKTYQVAHMKSKDRIRKLMNEILNGIKVLKLYAMELAFKDKVLAIRQBELKYLK 540
DB      481 AVNMAKTKTYQVAHMKSKDRIRKLMNEILNGIKVLKLYAMELAFKDKVLAIRQBELKYLK 540
QY      541 KSAYLAVGTFTWCTPFLVALCTFAVYVYIDENNILDAQTAFLVALFNILRPFNLILP 600
DB      541 KSAYLAVGTFTWCTPFLVALCTFAVYVYIDENNILDAQTAFLVALFNILRPFNLILP 600
QY      541 KSAYLAVGTFTWCTPFLVALCTFAVYVYIDENNILDAQTAFLVALFNILRPFNLILP 600
DB      541 KSAYLAVGTFTWCTPFLVALCTFAVYVYIDENNILDAQTAFLVALFNILRPFNLILP 600
QY      601 MVISIIQASVSLKRLAIFLSHELEPDSIERRVKQGGTNSITVNAATFTWARSDEPT 660
DB      601 MVISIIQASVSLKRLAIFLSHELEPDSIERRVKQGGTNSITVNAATFTWARSDEPT 660
QY      601 MVISIIQASVSLKRLAIFLSHELEPDSIERRVKQGGTNSITVNAATFTWARSDEPT 660
DB      601 MVISIIQASVSLKRLAIFLSHELEPDSIERRVKQGGTNSITVNAATFTWARSDEPT 660
QY      661 LMGITFSIPREGALVAVVQVQCGKSLSLALAMEDKYEGVAIKGSVAAYVPOQAWIQND 720
DB      661 LMGITFSIPREGALVAVVQVQCGKSLSLALAMEDKYEGVAIKGSVAAYVPOQAWIQND 720
QY      721 SLRENILFGQLEBPYRSYIQAACALLPDLIELPSGDRTEIGEXVNLSGQOKORVSLAR 780
DB      721 SLRENILFGQLEBPYRSYIQAACALLPDLIELPSGDRTEIGEXVNLSGQOKORVSLAR 780
QY      721 SLRENILFGQLEBPYRSYIQAACALLPDLIELPSGDRTEIGEXVNLSGQOKORVSLAR 780
DB      721 SLRENILFGQLEBPYRSYIQAACALLPDLIELPSGDRTEIGEXVNLSGQOKORVSLAR 780
QY      781 AVYSNADIIYLPDPLSAVDHVGKHIPEENVIGPKMLKNKRIILVTHSMSTLPQVDVYIV 840
DB      781 AVYSNADIIYLPDPLSAVDHVGKHIPEENVIGPKMLKNKRIILVTHSMSTLPQVDVYIV 840
QY      781 AVYSNADIIYLPDPLSAVDHVGKHIPEENVIGPKMLKNKRIILVTHSMSTLPQVDVYIV 840
DB      781 AVYSNADIIYLPDPLSAVDHVGKHIPEENVIGPKMLKNKRIILVTHSMSTLPQVDVYIV 840
QY      841 MSGGKISEMGSYQELLRDGAFAEFLRTYASTBOQDAEENGVTGVSQPGKEAKOMENGM 900
DB      841 MSGGKISEMGSYQELLRDGAFAEFLRTYASTBOQDAEENGVTGVSQPGKEAKOMENGM 900
QY      841 MSGGKISEMGSYQELLRDGAFAEFLRTYASTBOQDAEENGVTGVSQPGKEAKOMENGM 900
DB      841 MSGGKISEMGSYQELLRDGAFAEFLRTYASTBOQDAEENGVTGVSQPGKEAKOMENGM 900
QY      901 LVYDSAGKOLORQUSSSSYSGDISRRHNSYABLOKAEMAKKETWKLMEADKAOTGOYKL 960
DB      901 LVYDSAGKOLORQUSSSSYSGDISRRHNSYABLOKAEMAKKETWKLMEADKAOTGOYKL 960
QY      901 LVYDSAGKOLORQUSSSSYSGDISRRHNSYABLOKAEMAKKETWKLMEADKAOTGOYKL 960
DB      901 LVYDSAGKOLORQUSSSSYSGDISRRHNSYABLOKAEMAKKETWKLMEADKAOTGOYKL 960
QY      961 SVYWDYMKAIQGLFISFLSIFLPMCNHVSALASNWLSLMTDDPIVNGIQHTKRLSYYG 1020
DB      961 SVYWDYMKAIQGLFISFLSIFLPMCNHVSALASNWLSLMTDDPIVNGIQHTKRLSYYG 1020
QY      961 SVYWDYMKAIQGLFISFLSIFLPMCNHVSALASNWLSLMTDDPIVNGIQHTKRLSYYG 1020
DB      961 SVYWDYMKAIQGLFISFLSIFLPMCNHVSALASNWLSLMTDDPIVNGIQHTKRLSYYG 1020
QY      1021 ALGISOGIAVFGYMAVSIIGIILASRCIADVLHSIIRSPMSFPERIPSGULVNRFSKEL 1080
DB      1021 ALGISOGIAVFGYMAVSIIGIILASRCIADVLHSIIRSPMSFPERIPSGULVNRFSKEL 1080
QY      1021 ALGISOGIAVFGYMAVSIIGIILASRCIADVLHSIIRSPMSFPERIPSGULVNRFSKEL 1080
DB      1021 ALGISOGIAVFGYMAVSIIGIILASRCIADVLHSIIRSPMSFPERIPSGULVNRFSKEL 1080
QY      1081 DTVDMSIPEVIKMGMSLFNVIGACIVILATPIAIIIPPLGIYFFVQGFYVASSROL 1140
DB      1081 DTVDMSIPEVIKMGMSLFNVIGACIVILATPIAIIIPPLGIYFFVQGFYVASSROL 1140
QY      1081 DTVDMSIPEVIKMGMSLFNVIGACIVILATPIAIIIPPLGIYFFVQGFYVASSROL 1140
DB      1081 DTVDMSIPEVIKMGMSLFNVIGACIVILATPIAIIIPPLGIYFFVQGFYVASSROL 1140
QY      1141 KRLSVSRSPYVSHFNETLIGSVYIRAFEEQERFIHQSDLKVDENOKAYYSIVANRWLA 1200
DB      1141 KRLSVSRSPYVSHFNETLIGSVYIRAFEEQERFIHQSDLKVDENOKAYYSIVANRWLA 1200
QY      1141 KRLSVSRSPYVSHFNETLIGSVYIRAFEEQERFIHQSDLKVDENOKAYYSIVANRWLA 1200
DB      1141 KRLSVSRSPYVSHFNETLIGSVYIRAFEEQERFIHQSDLKVDENOKAYYSIVANRWLA 1200
QY      1201 VRLBCVNCIYLPALFAVISRHSLSAGLVGLSVYSIQVTTYVNLVRRSMSEMETNIVA 1260
DB      1201 VRLBCVNCIYLPALFAVISRHSLSAGLVGLSVYSIQVTTYVNLVRRSMSEMETNIVA 1260
QY      1201 VRLBCVNCIYLPALFAVISRHSLSAGLVGLSVYSIQVTTYVNLVRRSMSEMETNIVA 1260
DB      1201 VRLBCVNCIYLPALFAVISRHSLSAGLVGLSVYSIQVTTYVNLVRRSMSEMETNIVA 1260
QY      1261 VERLKEVSETEKEAPWQIOETAPPSSWPOVGRVEFRNYCLRYRBDLPVLRHINVTINGG 1320
DB      1261 VERLKEVSETEKEAPWQIOETAPPSSWPOVGRVEFRNYCLRYRBDLPVLRHINVTINGG 1320
QY      1261 VERLKEVSETEKEAPWQIOETAPPSSWPOVGRVEFRNYCLRYRBDLPVLRHINVTINGG 1320
DB      1261 VERLKEVSETEKEAPWQIOETAPPSSWPOVGRVEFRNYCLRYRBDLPVLRHINVTINGG 1320
QY      1321 EKVGIVGTGAKSSLTGLFRINESABGEIIIDGINIAKIGLHDLRFKTIITPODPYLF 1380
DB      1321 EKVGIVGTGAKSSLTGLFRINESABGEIIIDGINIAKIGLHDLRFKTIITPODPYLF 1380
QY      1321 EKVGIVGTGAKSSLTGLFRINESABGEIIIDGINIAKIGLHDLRFKTIITPODPYLF 1380
DB      1321 EKVGIVGTGAKSSLTGLFRINESABGEIIIDGINIAKIGLHDLRFKTIITPODPYLF 1380
QY      1381 SCSLRMNLDPFSQVSDERWTSLELAHKDVSALPDLDBECABGGNLSVGQQLVCL 1440
DB      1381 SCSLRMNLDPFSQVSDERWTSLELAHKDVSALPDLDBECABGGNLSVGQQLVCL 1440
QY      1381 SCSLRMNLDPFSQVSDERWTSLELAHKDVSALPDLDBECABGGNLSVGQQLVCL 1440
DB      1381 SCSLRMNLDPFSQVSDERWTSLELAHKDVSALPDLDBECABGGNLSVGQQLVCL 1440
QY      1441 ARALLRKTKIIVLDEATVAVDLETDLLIQSTIRFOFEDCTVLTIAHRLNTIMDYTRYIVL 1500
DB      1441 ARALLRKTKIIVLDEATVAVDLETDLLIQSTIRFOFEDCTVLTIAHRLNTIMDYTRYIVL 1500
QY      1441 ARALLRKTKIIVLDEATVAVDLETDLLIQSTIRFOFEDCTVLTIAHRLNTIMDYTRYIVL 1500
DB      1441 ARALLRKTKIIVLDEATVAVDLETDLLIQSTIRFOFEDCTVLTIAHRLNTIMDYTRYIVL 1500
QY      1501 DKGETIOEYGAQSDLLQQRGLFYMAKQAGLV 1531
DB      1501 DKGETIOEYGAQSDLLQQRGLFYMAKQAGLV 1531
```

Db 1501 DKGEIOEYGA5BDLQOGRGLFY5MAKDAGLV 1531

RESULT 4
US-08-463-179A-4
Sequence 4, Application US/08463179A
Patent No. 6001563
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Deele, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,179A
FILING DATE:

CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Deconti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: PQ1-002CP8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-179A-4

Query Match 80.7%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFSSAAGSDPLMNMNTTNTSNDFTYKCFONTLVVWPCYYLMAACPPHYLYISRH 60
Db 1 MALRGFSSAAGSDPLMNMNTTNTSNDFTYKCFONTLVVWPCYYLMAACPPHYLYISRH 60

QY 61 DRGVIQNTPLNKTALGFLIMTICWADLFYSFWERSRGIFLAVFVLSPTLIGITTLA 120
Db 61 DRGVIQNTPLNKTALGFLIMTICWADLFYSFWERSRGIFLAVFVLSPTLIGITTLA 120

QY 121 TPLIQERRRGVSSGIMLTFWLVAALVCAALILRSKIMTALKEBDAQVDLFRDIFYYFSS 180
Db 121 TPLIQERRRGVSSGIMLTFWLVAALVCAALILRSKIMTALKEBDAQVDLFRDIFYYFSS 180

QY 181 LLLIQVLVSGRSPSRPLFSETIHDNPNCPRESSASFLSRITFWMTGLIVNGYRPLEGSD 240
Db 181 LLLIQVLVSGRSPSRPLFSETIHDNPNCPRESSASFLSRITFWMTGLIVNGYRPLEGSD 240

QY 241 LMSLNKEDTSEQVAVLVKWKKECAKTRKQPVKVVYSSKDPQPKSSKVDANEVEAL 300

Db 241 LMSLNKEDTSEQVAVLVKWKKECAKTRKQPVKVVYSSKDPQPKSSKVDANEVEAL 300

QY 301 IVKSPQKSNP5LSEKVLKTPGPFLMSFPFKAIHDLMMSPQILKLLIKFVNDTKAPD 360
Db 301 IVKSPQKSNP5LSEKVLKTPGPFLMSFPFKAIHDLMMSPQILKLLIKFVNDTKAPD 360

QY 361 WQGFYTVLLFVTAQCLQTLVHOFVHICFVSGNRIKTAIVGAVRKALVITNSARKSTV 420
Db 361 WQGFYTVLLFVTAQCLQTLVHOFVHICFVSGNRIKTAIVGAVRKALVITNSARKSTV 420

QY 421 GEIVNLMSVDAQREMDLATYINMISAPLOVITLALYLMLNIGPSVLAVGAVMLAMPVN 480
Db 421 GEIVNLMSVDAQREMDLATYINMISAPLOVITLALYLMLNIGPSVLAVGAVMLAMPVN 480

QY 481 AVNANKTKTYQVAMKSKNRKIKLMEIINGIKVLKIYAMELAFKDYALAIROBELKVLK 540
Db 481 AVNANKTKTYQVAMKSKNRKIKLMEIINGIKVLKIYAMELAFKDYALAIROBELKVLK 540

QY 541 KSAVL5AVGPTWCTPELVALCTFAVYVITDENNILLDAQTAFSVSLAFNLRPNTLP 600
Db 541 KSAVL5AVGPTWCTPELVALCTFAVYVITDENNILLDAQTAFSVSLAFNLRPNTLP 600

QY 601 MVISSIVQASVSLKRLRIFLSHEELBPSIERRPVKDGGTNSITVBNATFTMARSDPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHEELBPSIERRPVKDGGTNSITVBNATFTMARSDPT 660

QY 661 LMGITFSPREGALVAVVGQVCGKSSLLSALLAEMDRVESHVAIKGSVAVYVQQA1QND 720
Db 661 LMGITFSPREGALVAVVGQVCGKSSLLSALLAEMDRVESHVAIKGSVAVYVQQA1QND 720

QY 721 SLRENILFGQLEBPYRSVIOACALLPDELIPSGRTERIGEGVNLSSGQKRVSLAR 780
Db 721 SLRENILFGQLEBPYRSVIOACALLPDELIPSGRTERIGEGVNLSSGQKRVSLAR 780

QY 781 AVYSNADIYFDDPLSAVDHVGKHPENYIGPKMKKNTRIIVTHSM5YLPOVDYIIV 840
Db 781 AVYSNADIYFDDPLSAVDHVGKHPENYIGPKMKKNTRIIVTHSM5YLPOVDYIIV 840

QY 841 MSGGKISEMGSYOELLARDGAFAPFLRTVASTBQDAEENGVTGVS6PGKEAKOMENG 900
Db 841 MSGGKISEMGSYOELLARDGAFAPFLRTVASTBQDAEENGVTGVS6PGKEAKOMENG 900

QY 901 LVNDSAGKQIQRQLSSSSSYSGDISRRHNSTAEIQKAEAKKEFWKLMEDKQOTGOVKL 960
Db 901 LVNDSAGKQIQRQLSSSSSYSGDISRRHNSTAEIQKAEAKKEFWKLMEDKQOTGOVKL 960

QY 961 SVYWDYKAIQLFISFLIFLPMCNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
Db 961 SVYWDYKAIQLFISFLIFLPMCNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020

QY 1021 ALGISQIAVFGYSMAVISGILLASRCLAHVDLHSLIRSPMSFBERP5GNLVNRF5KEL 1080
Db 1021 ALGISQIAVFGYSMAVISGILLASRCLAHVDLHSLIRSPMSFBERP5GNLVNRF5KEL 1080

QY 1081 DTVD5MTPEVIKMPM5SLFNVIACIVILATPAAIIPPLGLIFRFVORFYVASSROL 1140
Db 1081 DTVD5MTPEVIKMPM5SLFNVIACIVILATPAAIIPPLGLIFRFVORFYVASSROL 1140

QY 1141 KRLSVSRSPVYSHPNETLLGVSVIRAFEBQEPFIHOSDLKVBENQAVYPSIYANMWA 1200
Db 1141 KRLSVSRSPVYSHPNETLLGVSVIRAFEBQEPFIHOSDLKVBENQAVYPSIYANMWA 1200

QY 1201 VRL5CVN5CIVLPAALFAVISRSL5AGVLSV5YSLQVTTYLNMVLVR55SEMETIVA 1260
Db 1201 VRL5CVN5CIVLPAALFAVISRSL5AGVLSV5YSLQVTTYLNMVLVR55SEMETIVA 1260

QY 1261 VERLXE5SETEKAPMOIQETAPSSWPQVGRVFRNYCLARYEDLDVFLRHINVTNGG 1320
Db 1261 VERLXE5SETEKAPMOIQETAPSSWPQVGRVFRNYCLARYEDLDVFLRHINVTNGG 1320

QY 1321 EKVGIVGRTGAGSS5LTLGLFRIN5AEGRIIIDGINAIKIGLHDLRFKTTIIPQDPVLV 1380

Db 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEI1IDGINIAKIGLHDLRFKTIITIPDDPVLF 1380
Qy 1381 SGLSRMLNDPFSQYSDSEEWTSLELAHLKDFVSALPDKLDHECAGENGSLVGRQIVCL 1440
Db 1381 SGLSRMLNDPFSQYSDSEEWTSLELAHLKDFVSALPDKLDHECAGENGSLVGRQIVCL 1440
Qy 1441 ARALLRKTIIVLDEATPAVDLETFDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRIVL 1500
Db 1441 ARALLRKTIIVLDEATPAVDLETFDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRIVL 1500
Qy 1501 DKGEIOEGAPSDLLQORGLFYMAKXGIV 1531
Db 1501 DKGEIOEGAPSDLLQORGLFYMAKXGIV 1531

RESULT 5

US-08-461-384B-4

Sequence 4, Application US/08461384B
Patent No. 6025473

GENERAL INFORMATION:

APPLICANT: Cole, Susan P.C.

APPLICANT: Dealey, Roger G.

TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESSES:

ADDRESS: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS

STREET: Queen's University at Kingston

CITY: Kingston

STATE: Ontario

COUNTRY: CANADA

ZIP: K7L 3N6

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/461,384B

FILING DATE: 05-JUN-95

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 07/966,923

FILING DATE: 27-OCT-1992

APPLICATION NUMBER: 08/029,340

FILING DATE: 8-MAR-1993

APPLICATION NUMBER: 08/141,893

FILING DATE: 26-OCT-1993

APPLICATION NUMBER: 08/407,207

FILING DATE: 20-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Steeg, Carol Miernicki

REGISTRATION NUMBER: 39,539

REFERENCE/DOCKET NUMBER: 01547

TELECOMMUNICATION INFORMATION:

TELEPHONE: (613) 545-2342

TELEFAX: (613) 545-6853

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1531 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-461-384B-4

Query Match 80.7%; Score 7860; DB 2; Length 1531;

Best local Similarity 100.0%; Pred. No. 0;

Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRFGCSADGSDPLMDNMVNTSNPDFTKCFONTVLVWVPCFYLMACFPFYFLYLSRH 60
Db 1 MALRFGCSADGSDPLMDNMVNTSNPDFTKCFONTVLVWVPCFYLMACFPFYFLYLSRH 60
Qy 61 DRGIYQMTPLNKTALGFLLMTVCMADLFYSPWERSRGIFLAPVPLVSPTLIGITTLIA 120
Db 61 DRGIYQMTPLNKTALGFLLMTVCMADLFYSPWERSRGIFLAPVPLVSPTLIGITTLIA 120

Db 61 DRGIYQMTPLNKTALGFLLMTVCMADLFYSPWERSRGIFLAPVPLVSPTLIGITTLIA 120
Qy 121 TFLIOLERRKVOSSGIMLTFMLVALYCALAIIKSKIMTALKEDAQVDLFEDITFYVYFS 180
Db 121 TFLIOLERRKVOSSGIMLTFMLVALYCALAIIKSKIMTALKEDAQVDLFEDITFYVYFS 180
Qy 181 LLLIQLVLSGCSDBSPLESETIHDNCPRESSASFSLRITFWMTTGLIVRGYRQPLEBSD 240
Db 181 LLLIQLVLSGCSDBSPLESETIHDNCPRESSASFSLRITFWMTTGLIVRGYRQPLEBSD 240
Qy 241 LMSLNKEDTSQVVPVLYKMKKECAKTRKQPVVYVSSSKDPAOPKSSSKDANEVEVAL 300
Db 241 LMSLNKEDTSQVVPVLYKMKKECAKTRKQPVVYVSSSKDPAOPKSSSKDANEVEVAL 300
Qy 301 IVKSPQKEMNPSLFKVLKTFGPYFLMSFPFKAHDMFMESGQILKLIKFVNDTKAPD 360
Db 301 IVKSPQKEMNPSLFKVLKTFGPYFLMSFPFKAHDMFMESGQILKLIKFVNDTKAPD 360
Qy 361 WQGYFYTVLLFVTKCLQTLVLIHQYFHICFVSGMKTAVIGAVYRKALVITNSARKSTV 420
Db 361 WQGYFYTVLLFVTKCLQTLVLIHQYFHICFVSGMKTAVIGAVYRKALVITNSARKSTV 420
Qy 421 GEIVNIMSVDQRPMDLATYINMIMSAPLOVITLALYLMNLGPSVLGAVAVMYLMPVYN 480
Db 421 GEIVNIMSVDQRPMDLATYINMIMSAPLOVITLALYLMNLGPSVLGAVAVMYLMPVYN 480
Qy 481 AYMAKTKTYQVAHMKSKDNRIKLMNBIINGIKVLKUYAMELAFKDYALAIROBELKYLK 540
Db 481 AYMAKTKTYQVAHMKSKDNRIKLMNBIINGIKVLKUYAMELAFKDYALAIROBELKYLK 540
Qy 541 KSAVLSAVGTTTWTCTPFLVALCTFPAYVYTIIDENNIIIDAOCTAPFSLAFNLRPPNLI 600
Db 541 KSAVLSAVGTTTWTCTPFLVALCTFPAYVYTIIDENNIIIDAOCTAPFSLAFNLRPPNLI 600
Qy 601 MVISIVQASVSLKRIKFLSHEELBPSIRRPVKQGGNISTVRATFTMARSDPT 660
Db 601 MVISIVQASVSLKRIKFLSHEELBPSIRRPVKQGGNISTVRATFTMARSDPT 660
Qy 661 LMGITFSPREGALVAVVQVCGKSSLSALLAEMDKVGHVAIKGSVAVYVPOQAMIOND 720
Db 661 LMGITFSPREGALVAVVQVCGKSSLSALLAEMDKVGHVAIKGSVAVYVPOQAMIOND 720
Qy 721 SLRENIIFGCOLLEBPYRSVIOACALLPDLIELPSGDRTEIGEKGVNISGQOKORVSLAR 780
Db 721 SLRENIIFGCOLLEBPYRSVIOACALLPDLIELPSGDRTEIGEKGVNISGQOKORVSLAR 780
Qy 781 AVYSNADIVLEDDPLSADAHVGHIFENYVIGPKMLKNKTRILVTHSMSTLPOVDVITV 840
Db 781 AVYSNADIVLEDDPLSADAHVGHIFENYVIGPKMLKNKTRILVTHSMSTLPOVDVITV 840
Qy 841 MSGKISEMGSYOBELIARDGAFAEFLRTYASTEOBDAEENGVTGVSQPGKEAKOMENG 900
Db 841 MSGKISEMGSYOBELIARDGAFAEFLRTYASTEOBDAEENGVTGVSQPGKEAKOMENG 900
Qy 901 LVYDSAGKOLOROLSSSSSYSGDISRRHNSYAELOKAPAKKEETWKLMEADKAQTGYKL 960
Db 901 LVYDSAGKOLOROLSSSSSYSGDISRRHNSYAELOKAPAKKEETWKLMEADKAQTGYKL 960
Qy 961 SVYWDYMAKIGLFTSFLSIFLMCHNVASALSNWLSLMTDDPIYNGOETHKTVLSYVG 1020
Db 961 SVYWDYMAKIGLFTSFLSIFLMCHNVASALSNWLSLMTDDPIYNGOETHKTVLSYVG 1020
Qy 1021 ALGISGIAVFGYSMAVSIIGLILASRCHVDLHSILSPMSFFERTSGNULVNFSEKL 1080
Db 1021 ALGISGIAVFGYSMAVSIIGLILASRCHVDLHSILSPMSFFERTSGNULVNFSEKL 1080
Qy 1081 DTVDSMIDEVIMKFMGSLFNVIAGCTVILATPAAIIIPVLGIYFFVQRPYVASSQOL 1140
Db 1081 DTVDSMIDEVIMKFMGSLFNVIAGCTVILATPAAIIIPVLGIYFFVQRPYVASSQOL 1140
Qy 1141 KRLESVSSPYVSHNETLLGVSVIRAFEEBERFHOSDLVDENOKAYYSIVANRWLA 1200
Db 1141 KRLESVSSPYVSHNETLLGVSVIRAFEEBERFHOSDLVDENOKAYYSIVANRWLA 1200

QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGVSYSLOVTTYLANMLVRMSSEMETNIVA 1260
Db 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGVSYSLOVTTYLANMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFRNVCYRREDLDFVLRHINTVINGG 1320
Db 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFRNVCYRREDLDFVLRHINTVINGG 1320
QY 1321 EKVGIQVHTGAGKSLTGLFRINESAGEIITIDGINAKIGLHDLRFKTIITIPQDPVLF 1380
Db 1321 EKVGIQVHTGAGKSLTGLFRINESAGEIITIDGINAKIGLHDLRFKTIITIPQDPVLF 1380
QY 1381 SGSLRMLNDPFSQVSDSEEWTSLSLAHLKDFVSALPDGLDHECAGGGENLSVGOQOLVCL 1440
Db 1381 SGSLRMLNDPFSQVSDSEEWTSLSLAHLKDFVSALPDGLDHECAGGGENLSVGOQOLVCL 1440
QY 1441 ABALLRKTILVLDEATAVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVYL 1500
Db 1441 ABALLRKTILVLDEATAVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVYL 1500
QY 1501 DKGEIOEYGA PSDLLQORGLFYSMACKAGLV 1531
Db 1501 DKGEIOEYGA PSDLLQORGLFYSMACKAGLV 1531

RESULT 6

US-09-647-140B-19
; Sequence 19, Application US/09647140B
; Patent No. 6803184
; GENERAL INFORMATION:
; APPLICANT: Fox Chase Cancer Center
; APPLICANT: Kruth, Gary D.
; APPLICANT: Lee, Kun
; APPLICANT: Belinsky, Martin G.
; APPLICANT: Bain, Lisa J.
; TITLE OF INVENTION: MRP-Related ABC Transporter Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use Thereof
; FILE REFERENCE: PCCC 98-02
; CURRENT APPLICATION NUMBER: US/09/647, 140B
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US99/06644
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/079, 759
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/095, 153
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-647-140B-19

Query Match 80.7%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pctd. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTVNTSNPDFTKCFONTVLVWPCFYLMACPFYLYLSRH 60
Db 1 MALRGFCSADGSDPLMDMNTVNTSNPDFTKCFONTVLVWPCFYLMACPFYLYLSRH 60
QY 61 DRGVIQMTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLAVFLVSPITLLGITLLA 120
Db 61 DRGVIQMTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLAVFLVSPITLLGITLLA 120
QY 121 TPLIQLERRKGVSSGIMLFEMVLVAVCALAILBSKIMTALXBAQVDFRDTTFYVYS 180
Db 121 TPLIQLERRKGVSSGIMLFEMVLVAVCALAILBSKIMTALXBAQVDFRDTTFYVYS 180
QY 181 LLLIQVLVSCFSDRSPLFSETIHDPNCPBSSASFLSRITFWWITGLIVRGYRQPLEGSD 240
Db 181 LLLIQVLVSCFSDRSPLFSETIHDPNCPBSSASFLSRITFWWITGLIVRGYRQPLEGSD 240

Db 181 LLLIQVLVSCFSDRSPLFSETIHDPNCPBSSASFLSRITFWWITGLIVRGYRQPLEGSD 240
QY 241 LMSLNKEDTSGQVTVLVKWKKECATRKPQVYVYSSKDPAPKSSKTDANEVEAL 300
Db 241 LMSLNKEDTSGQVTVLVKWKKECATRKPQVYVYSSKDPAPKSSKTDANEVEAL 300
QY 301 IVKSPQKMNPSLKVLYKTPGPYELMSFPFKALHDMFSGPOLIKLLIKFVNDTPAPD 360
Db 301 IVKSPQKMNPSLKVLYKTPGPYELMSFPFKALHDMFSGPOLIKLLIKFVNDTPAPD 360
QY 361 WQGYFYTVLLFVTAQOTLVHLQYFHIQFVSGMRKTAVIGAVVRKALVITNSARKSTV 420
Db 361 WQGYFYTVLLFVTAQOTLVHLQYFHIQFVSGMRKTAVIGAVVRKALVITNSARKSTV 420
QY 421 GEIVNLSVDAQRFMDLATTYINMIWSAPLOYITALYLLMTNLGPSVLAGAVMYLWVFN 480
Db 421 GEIVNLSVDAQRFMDLATTYINMIWSAPLOYITALYLLMTNLGPSVLAGAVMYLWVFN 480
QY 481 AVMAKTKTYOVAMKSKDNRIKLMNEILNGIKVLYAMELAKDVKVLAIROBELKVLK 540
Db 481 AVMAKTKTYOVAMKSKDNRIKLMNEILNGIKVLYAMELAKDVKVLAIROBELKVLK 540
QY 541 KSAVLASVGTFTWCTPFLVALCTFAVYVITIDENNIIDAQTAFFVSLALFNILRPPLNLP 600
Db 541 KSAVLASVGTFTWCTPFLVALCTFAVYVITIDENNIIDAQTAFFVSLALFNILRPPLNLP 600
QY 601 MVISSIVQASVLSKRLRIFLSHEELBPSIERRPVKDGGGTSITVNAATFTWARSDEPT 660
Db 601 MVISSIVQASVLSKRLRIFLSHEELBPSIERRPVKDGGGTSITVNAATFTWARSDEPT 660
QY 661 LMGITFSPBGAALVAVVQVCGCKSSLSLALLAEMDVBEHVAKGSVAVYPOQAWIQND 720
Db 661 LMGITFSPBGAALVAVVQVCGCKSSLSLALLAEMDVBEHVAKGSVAVYPOQAWIQND 720
QY 721 SLRENILFGQLSEPPYRSVIOACALLPDLEILPSEGRTEIGEGVNLSSGQKORVSLAR 780
Db 721 SLRENILFGQLSEPPYRSVIOACALLPDLEILPSEGRTEIGEGVNLSSGQKORVSLAR 780
QY 781 AYSNADIYLPDPLSAVDAHVGHKIPENYIGPKMLKNKTRILLVTHSMSYLPOVDYIIV 840
Db 781 AYSNADIYLPDPLSAVDAHVGHKIPENYIGPKMLKNKTRILLVTHSMSYLPOVDYIIV 840
QY 841 MSGGKISEMSYQELLARDAPAEFLRTASTQEOEOAENGVTVSGSPKKAQOMNGM 900
Db 841 MSGGKISEMSYQELLARDAPAEFLRTASTQEOEOAENGVTVSGSPKKAQOMNGM 900
QY 901 LVYDSAGKQOROLSSSSSYSGDISRHHNSTAELOKAEKKEETWKLMEADKXQOTGVKL 960
Db 901 LVYDSAGKQOROLSSSSSYSGDISRHHNSTAELOKAEKKEETWKLMEADKXQOTGVKL 960
QY 961 SYVWDYMKALGLFISPLSIFLPMCNHVSALASNYMLSLWTDPIVNGTOBHTKVRLSVYG 1020
Db 961 SYVWDYMKALGLFISPLSIFLPMCNHVSALASNYMLSLWTDPIVNGTOBHTKVRLSVYG 1020
QY 1021 ALGISQIAVFGYSMAVSTIGILLASRCLHVDLHSLIRSPMSFEFRPNSGLNVRFSKEL 1080
Db 1021 ALGISQIAVFGYSMAVSTIGILLASRCLHVDLHSLIRSPMSFEFRPNSGLNVRFSKEL 1080
QY 1081 DTVDSMIPETIKPMFMSLFVIGACIVILLATPIAIIIPPLGIYFVFVORFYVASSROL 1140
Db 1081 DTVDSMIPETIKPMFMSLFVIGACIVILLATPIAIIIPPLGIYFVFVORFYVASSROL 1140
QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200
Db 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENOKAYYPSIVANRWLA 1200
QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGVSYSLOVTTYLANMLVRMSSEMETNIVA 1260
Db 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGVSYSLOVTTYLANMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFRNVCYRREDLDFVLRHINTVINGG 1320
Db 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFRNVCYRREDLDFVLRHINTVINGG 1320

QY 1321 EKVGIVGRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLARFKITIIIPQDPVLF 1380
DB 1321 EKVGIVGRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLARFKITIIIPQDPVLF 1380
QY 1381 SGSLRMLNDPFSQYSDERWMTSLBLAHLKDFVSALPDLKDHECAEGENLSVGQRLVCL 1440
DB 1381 SGSLRMLNDPFSQYSDERWMTSLBLAHLKDFVSALPDLKDHECAEGENLSVGQRLVCL 1440
QY 1441 ARALARKTILVLEATAVADLETDDLIQSTIRIQFEDCTVLTIAHRLNTIMDTYRVTL 1500
DB 1441 ARALARKTILVLEATAVADLETDDLIQSTIRIQFEDCTVLTIAHRLNTIMDTYRVTL 1500
QY 1501 DKGEIOEYGA PSDLLQORGLFYMAKADAGLV 1531
DB 1501 DKGEIOEYGA PSDLLQORGLFYMAKADAGLV 1531

RESULT 7
US-08-141-893-2
Sequence 2, Application US/08141893
Patent No. 5489519
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEIN
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 STATE STREET, SUITE 510
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/141,893
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923; 08/029,340
FILING DATE: 27-OCT-1992; 8-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: POI-002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5149
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-141-893-2

Query Match 80.6%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCAGSDGDPIDMNMVNTSNPDFTKCFONTVLVWPCFYIMACFPFYLYLSRH 60
DB 1 MALRGFCAGSDGDPIDMNMVNTSNPDFTKCFONTVLVWPCFYIMACFPFYLYLSRH 60
QY 61 DRGYQMPPLNKTALGELMLIVCMADLFYFWERSRGIETLAPVETVSLGTTTLA 120
DB 61 DRGYQMPPLNKTALGELMLIVCMADLFYFWERSRGIETLAPVETVSLGTTTLA 120

QY 121 TFLIQERRRGVSSGIMLTFWLVALVCAALILRSKIMTLAKEDAOVDLFEDITFYFFS 180
DB 121 TFLIQERRRGVSSGIMLTFWLVALVCAALILRSKIMTLAKEDAOVDLFEDITFYFFS 180
QY 181 LLLIQVLSCGSDSPLEFSETIHDNCPSSASFLSRITFMWTTGLIVGGRPLGSD 240
DB 181 LLLIQVLSCGSDSPLEFSETIHDNCPSSASFLSRITFMWTTGLIVGGRPLGSD 240
QY 241 LMSLNKEDTSEQVVPVVKWKKECAKTRKQPVVYVSSKDPAOKESSKYDADEVEAL 300
DB 241 LMSLNKEDTSEQVVPVVKWKKECAKTRKQPVVYVSSKDPAOKESSKYDADEVEAL 300
QY 301 IVKSPQKEMNPSLEKVLKYTFGPFLMSPEFKAHDLMMFSGPOLKLLIKFVNDTKAPD 360
DB 301 IVKSPQKEMNPSLEKVLKYTFGPFLMSPEFKAHDLMMFSGPOLKLLIKFVNDTKAPD 360
QY 361 WQGYFYTVLLEPVTCLOTIVLHOYFHIQVSGMKIKTRAVGAVYRKALVITNSARKSTV 420
DB 361 WQGYFYTVLLEPVTCLOTIVLHOYFHIQVSGMKIKTRAVGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINMISAPLQVITLALYLMNLGSPVLGAVAVMVLMPVN 480
DB 421 GEIVNLSVDAQRFMDLATYINMISAPLQVITLALYLMNLGSPVLGAVAVMVLMPVN 480
QY 481 AVMAKTKTYOVAMKSKDNRIKLMETLNGIKVLKIYAMELAKDKYLAIROBELKYLK 540
DB 481 AVMAKTKTYOVAMKSKDNRIKLMETLNGIKVLKIYAMELAKDKYLAIROBELKYLK 540
QY 541 KSAVLSAVGFTWCTPFLVALCTFAVYVITDENNIIDAOFAFSLFNLRLPINTLP 600
DB 541 KSAVLSAVGFTWCTPFLVALCTFAVYVITDENNIIDAOFAFSLFNLRLPINTLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHELEPDSIERRPVKOGGTNSITVNAFTWASDPPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHELEPDSIERRPVKOGGTNSITVNAFTWASDPPT 660
QY 661 LMGITFSPBGALVAVVQVCGGSSLSALLAEMDKYEGVALIKGSYAYVPOQAMIOND 720
DB 661 LMGITFSPBGALVAVVQVCGGSSLSALLAEMDKYEGVALIKGSYAYVPOQAMIOND 720
QY 721 SLRENILPGCOLERPYSRVIOACALLPDEILSGDRTLEGKGVNLSGGOKRVSILAR 780
DB 721 SLRENILPGCOLERPYSRVIOACALLPDEILSGDRTLEGKGVNLSGGOKRVSILAR 780
QY 781 AVYSNADITLDDPPLSAVDAHVGHIFENYIGPKGMLKNKRIIVTHSMSTLPQVDVIV 840
DB 781 AVYSNADITLDDPPLSAVDAHVGHIFENYIGPKGMLKNKRIIVTHSMSTLPQVDVIV 840
QY 841 MSGKISEMSGYOELIARDGAFAPFLRTYASTEQDQAEENGVTGVSQGRKAKOMENG 900
DB 841 MSGKISEMSGYOELIARDGAFAPFLRTYASTEQDQAEENGVTGVSQGRKAKOMENG 900
QY 901 LVTSBAGKOLOROSSSSSYSGDISRHNSYAELOKAEKKEFTWKMEADKAQCGVYL 960
DB 901 LVTSBAGKOLOROSSSSSYSGDISRHNSYAELOKAEKKEFTWKMEADKAQCGVYL 960
QY 961 SVYMDYKAIGLFTSIFLFLMCNHSALSNYWLMTDPTVNGQETHKRLSYG 1020
DB 961 SVYMDYKAIGLFTSIFLFLMCNHSALSNYWLMTDPTVNGQETHKRLSYG 1020
QY 1021 ALGISQGIADVGSMAVSIIGIILASRCLHVDLHLSILSPMSPEFTPSGMLVNFSEKL 1080
DB 1021 ALGISQGIADVGSMAVSIIGIILASRCLHVDLHLSILSPMSPEFTPSGMLVNFSEKL 1080
QY 1081 DTUDSMIEPVKMPKGSIFNVIGACIVYLLATPTIAIIIPGLGLYFVQRFYVASSQOL 1140
DB 1081 DTUDSMIEPVKMPKGSIFNVIGACIVYLLATPTIAIIIPGLGLYFVQRFYVASSQOL 1140
QY 1141 KRLBSVRSPPYSHNETLIGSVIRAFEBQERFIHOSDLKYDENOKAYYSIVANRMLA 1200
DB 1141 KRLBSVRSPPYSHNETLIGSVIRAFEBQERFIHOSDLKYDENOKAYYSIVANRMLA 1200
QY 1201 VRLBCVGNICIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVWMSSEMTNIVA 1260

Db 1201 VRLCEVGNCTVLFALFAVISRHSLSAGLVGSYSLSQVTTYINMLVRMSSEMETNIVA 1260
Qy 1261 VERLKEVSETEKEAPWQIQETAPPSWPOQVGRVERRANCYARREDLVLRHINVTNGG 1320
Db 1261 VERLKEVSETEKEAPWQIQETRPPSPQVORVERRNCCLYRBDLVLRHINVTNGG 1320
Qy 1321 EKVGVGTGAGKSLTGLFRINSEAGEIIIDGINAKIGLHDLRFKTIIPQDPVL 1380
Db 1321 EKVGVGTGAGKSLTGLFRINSEAGEIIIDGINAKIGLHDLRFKTIIPQDPVL 1380
Qy 1381 SGSLRMLNDPFSQYSDSEWMTSLBLAKDFFVSALPDKLDHCAEGGENLSVGQRQVCL 1440
Db 1381 SGSLRMLNDPFSQYSDSEWMTSLBLAKDFFVSALPDKLDHCAEGGENLSVGQRQVCL 1440
Qy 1441 ARALLRKTILVDEATAVDLETDLIQSTIRTOFEDCTVLTAAHRLNTMDYTRVVL 1500
Db 1441 ARALLRKTILVDEATAVDLETDLIQSTIRTOFEDCTVLTAAHRLNTMDYTRVVL 1500
Qy 1501 DKGEIOEYGAAPSDLLQORGLFYMAKQAGLV 1531
Db 1501 DKGEIOEYGAAPSDLLQORGLFYMAKQAGLV 1531

RESULT 8
US-08-463-0928-2
Sequence 2, Application US/084630928
Patent No. 5766880
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,0928
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: Q1546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-0928-2

Query Match 80.6%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MALRGFCSAGSDPLDMANVTNNTSNDFPKCFONTLVWVPCYLWACPFYLYSRH 60
Db 1 MALRGFCSAGSDPLDMANVTNNTSNDFPKCFONTLVWVPCYLWACPFYLYSRH 60
Qy 61 DRGYIQMTPLNKTALGFLMIYCMADLFYSFWERSRGIFLAPVFLVSPTLGITTLA 120
Db 61 DRGYIQMTPLNKTALGFLMIYCMADLFYSFWERSRGIFLAPVFLVSPTLGITTLA 120
Qy 121 TFLIQLERRKGVQSSGIMLTFMLVYCALAILRSKIMTALKEBAQVDFRDTFFYYFS 180
Db 121 TFLIQLERRKGVQSSGIMLTFMLVYCALAILRSKIMTALKEBAQVDFRDTFFYYFS 180
Qy 181 LLLIQVLSCFSDPSLFEFTIHDPNCPBSSASFSLRITFMWITGLIVRGYRPLGSD 240
Db 181 LLLIQVLSCFSDPSLFEFTIHDPNCPBSSASFSLRITFMWITGLIVRGYRPLGSD 240
Qy 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAQPKESKVDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSKDPAQPKESKVDANEVEAL 300
Qy 301 IVSPQKEMNPSLFKVLKTFGEYFLMSFPFKAHIDLMFSGPQILKLIKPVNDTAPD 360
Db 301 IVSPQKEMNPSLFKVLKTFGEYFLMSFPFKAHIDLMFSGPQILKLIKPVNDTAPD 360
Qy 361 WQGYFYTVLLFVTAQOTLVHQQYFHTCFSGMRKTAIVGAYYRKALVTNSRKSTV 420
Db 361 WQGYFYTVLLFVTAQOTLVHQQYFHTCFSGMRKTAIVGAYYRKALVTNSRKSTV 420
Qy 421 GEIVNLMSVDQRFMDLATYINMIWSAPLOVIALYLMNLGSPVLAVVWVLPVN 480
Db 421 GEIVNLMSVDQRFMDLATYINMIWSAPLOVIALYLMNLGSPVLAVVWVLPVN 480
Qy 481 AVNAAKTKTYQVAHMSKDNRIKLMNELLNGIKVLYLWELAFKDKVLAIRBELKVLK 540
Db 481 AVNAAKTKTYQVAHMSKDNRIKLMNELLNGIKVLYLWELAFKDKVLAIRBELKVLK 540
Qy 541 KSAVYLSAVGFTWVCTPELVALCTPAVYVITIDNNILDAQTAVSLAFNILLRPNILP 600
Db 541 KSAVYLSAVGFTWVCTPELVALCTPAVYVITIDNNILDAQTAVSLAFNILLRPNILP 600
Qy 601 MVYSSIVQASVSLKRLRIFLSHELEBPSIERRPVKDGSGTNSITYVNAFTWASDPPT 660
Db 601 MVYSSIVQASVSLKRLRIFLSHELEBPSIERRPVKDGSGTNSITYVNAFTWASDPPT 660
Qy 661 LMGITFSIPGALVAVVGVGCGKSLSLALAEMLVEBGHVAIKGSVAVVPQAWIOND 720
Db 661 LMGITFSIPGALVAVVGVGCGKSLSLALAEMLVEBGHVAIKGSVAVVPQAWIOND 720
Qy 721 SLAENILFGQLEBPYRSTYQACALLPDLIELPSGDRTEIGKGVNLSSGQQRVSILAR 780
Db 721 SLAENILFGQLEBPYRSTYQACALLPDLIELPSGDRTEIGKGVNLSSGQQRVSILAR 780
Qy 781 AVYSNADIVLPDPLSAVDAAVGHKIFENVIGKGMKNKTRILVTHSMYSYLCQVDYIV 840
Db 781 AVYSNADIVLPDPLSAVDAAVGHKIFENVIGKGMKNKTRILVTHSMYSYLCQVDYIV 840
Qy 841 MSGKTISEMSYQBLARDGAFELFTYASTQEOBDAENGTYGSGPKEAKOMENGM 900
Db 841 MSGKTISEMSYQBLARDGAFELFTYASTQEOBDAENGTYGSGPKEAKOMENGM 900
Qy 901 LVYDSAGKQIQRLSSSSVSQDISRHHSNTAEIQRAEAKKEFTWKLMEADKXQGTQVKL 960

```

Db 901 LVTSAGQLOKROUSSSSSSYSGDISRHNSIAELQKAKKEETKMEAKQOQGYKL 960
Qy 961 SVYDYMKAIGLFTISFLIFLMCMHVSALASNYLWLSMTDDPIVNGTOEHTKYRLSYG 1020
Db 961 SVYDYMKAIGLFTISFLIFLMCMHVSALASNYLWLSMTDDPIVNGTOEHTKYRLSYG 1020
Qy 1021 ALGISQGIAGVGYSAVNSIGSILASRCLHVDLHLSILSPMSFPERTPSGNLVNRFSKEL 1080
Db 1021 ALGISQGIAGVGYSAVNSIGSILASRCLHVDLHLSILSPMSFPERTPSGNLVNRFSKEL 1080
Qy 1081 DTUOSMTPEVIMKMGSLFNVIAGACIVILLATPIAIIIPPLGLIFFVQRFYVASSQOL 1140
Db 1081 DTUOSMTPEVIMKMGSLFNVIAGACIVILLATPIAIIIPPLGLIFFVQRFYVASSQOL 1140
Qy 1141 KRLSVSRSPYSHFNETLIGSVIRAFEEQERFIHQSDLKVDENQKAYYSIVANRWLA 1200
Db 1141 KRLSVSRSPYSHFNETLIGSVIRAFEEQERFIHQSDLKVDENQKAYYSIVANRWLA 1200
Qy 1201 VRLBCVAGNCIVLPAALFPAVISRHSISAGLVGLSVSYSLQVTTYLNMVLVMSSEMETNIVA 1260
Db 1201 VRLBCVAGNCIVLPAALFPAVISRHSISAGLVGLSVSYSLQVTTYLNMVLVMSSEMETNIVA 1260
Qy 1261 VERLKEVETEKEAPWQIOETAPPSWPQVGRVERFRNYCLAREDDLPVLRHINVTNGG 1320
Db 1261 VERLKEVETEKEAPWQIOETAPPSWPQVGRVERFRNYCLAREDDLPVLRHINVTNGG 1320
Qy 1321 EKVGIVGTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPODPLVF 1380
Db 1321 EKVGIVGTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPODPLVF 1380
Qy 1381 SGLSRMNLDPFSQYSDEEWTSLBELAHLKDPVSALPDKLDHCAEGENLSVGQOLVCL 1440
Db 1381 SGLSRMNLDPFSQYSDEEWTSLBELAHLKDPVSALPDKLDHCAEGENLSVGQOLVCL 1440
Qy 1441 ARALLRKTIIVLDBATAVVDLETDLIQSTIRTPFECTVLTANRANTIMDYTRIVYL 1500
Db 1441 ARALLRKTIIVLDBATAVVDLETDLIQSTIRTPFECTVLTANRANTIMDYTRIVYL 1500
Qy 1501 DKGEIOEYGAPSDILLQORGLFYSMAKDAGLV 1531
Db 1501 DKGEIOEYGAPSDILLQORGLFYSMAKDAGLV 1531

```

RESULT 9

US-08-462-109A-2
Sequence 2, Application US/08462109A

GENERAL INFORMATION:

APPLICANT: Cole, Susan P. C.
APPLICANT: Dealey, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING
TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS
NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESS: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/462,109A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/966,923

FILING DATE: 27-OCT-1992

```

APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: PQI-002CP4
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-2

```

Query Match 80.6%; Score 7849; DB 1; Length 1531;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy 1 MALRGCSADGSDPLMDMNTWNTSNPDFTKCFQNTVLVWPCHYLMACPEFFLYLSRH 60
Db 1 MALRGCSADGSDPLMDMNTWNTSNPDFTKCFQNTVLVWPCHYLMACPEFFLYLSRH 60
Qy 61 DRGIQNTPLNKTALGFLIMTYCWDLPFSFWERSRGIFLAVPLVSPTLGITTLLA 120
Db 61 DRGIQNTPLNKTALGFLIMTYCWDLPFSFWERSRGIFLAVPLVSPTLGITTLLA 120
Qy 121 TFLQLRRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEBDAQVDLPEDITFYFFS 180
Db 121 TFLQLRRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEBDAQVDLPEDITFYFFS 180
Qy 181 LLLIQVLVSCFSDRSLPFSETHIDNPNCPESASFLSRITFWMTGLIVRGYRQPLEGSD 240
Db 181 LLLIQVLVSCFSDRSLPFSETHIDNPNCPESASFLSRITFWMTGLIVRGYRQPLEGSD 240
Qy 241 LMSLNKEDTSQVIVLVKMKKCACTRKQPVVYVSSSDPADPKSSSKYDANEVYAL 300
Db 241 LMSLNKEDTSQVIVLVKMKKCACTRKQPVVYVSSSDPADPKSSSKYDANEVYAL 300
Qy 301 IVKSPKEMNPSLKVLYKTGPYFLMSFPFKAIHDLMMFSGPOLIKLIFVNDTKAPD 360
Db 301 IVKSPKEMNPSLKVLYKTGPYFLMSFPFKAIHDLMMFSGPOLIKLIFVNDTKAPD 360
Qy 361 WQGYFTVLVFTVACLOTVLHQYFHIQVSGMRKTAIGAVYRKALVITNSARKSTV 420
Db 361 WQGYFTVLVFTVACLOTVLHQYFHIQVSGMRKTAIGAVYRKALVITNSARKSTV 420
Qy 421 GEIVNLSVDAQRMIDLATTYINMTWSAPLOYITALYLLMLNLSGSYLAGVAVMYLMPVN 480
Db 421 GEIVNLSVDAQRMIDLATTYINMTWSAPLOYITALYLLMLNLSGSYLAGVAVMYLMPVN 480
Qy 481 AVMAKTKTYOVAMHKSNDNRKIKLNNELINGIKYLKLYAMELAKPDVYLAIRBELKYLK 540
Db 481 AVMAKTKTYOVAMHKSNDNRKIKLNNELINGIKYLKLYAMELAKPDVYLAIRBELKYLK 540
Qy 541 KSAVLSAVGFTWCTPFLVALCTFAVYVITDENNIIDAQTAFFSLLFNILRPPLNLP 600
Db 541 KSAVLSAVGFTWCTPFLVALCTFAVYVITDENNIIDAQTAFFSLLFNILRPPLNLP 600
Qy 601 MVISSIVQSVSLKRLRIFLSHEBELPDSIERRPVKDGGGNSITVNRATPTMARSDPT 660
Db 601 MVISSIVQSVSLKRLRIFLSHEBELPDSIERRPVKDGGGNSITVNRATPTMARSDPT 660
Qy 661 LMGITFSTPEGALVAVVGQVCGKSLSLALLAEMDVVEGHVATKGSAAVYPOQAMTOND 720
Db 661 LMGITFSTPEGALVAVVGQVCGKSLSLALLAEMDVVEGHVATKGSAAVYPOQAMTOND 720

```

[illegible]

```

? ZFP: K7L 3N6
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: ASCII text
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/460,907B
? FILING DATE: 05-JUN-1995
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/966,923
? FILING DATE: 27-OCT-1992
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/029,340
? FILING DATE: 8-MAR-1993
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/141,893
? FILING DATE: 26-OCT-1993
? CLASSIFICATION: 424
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/407,207
? FILING DATE: 20-MAR-1995
? CLASSIFICATION: 424
? ATTORNEY/AGENT INFORMATION:
? NAME: Steeg, Carol Miernicki
? REGISTRATION NUMBER: 39,539
? REFERENCE/DOCKET NUMBER: 01551
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (613) 545-2342
? TELEFAX: (613) 545-6853
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1531 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-460-907B-2
?
Query Match 80.6%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0.
?
QY 1 MALRGFSCADSGDPLDMNVTWNTSNPDFTKCFONTVLVWVPCFYLMACPEFFLYLSRH 60
DB 1 MALRGFSCADSGDPLDMNVTWNTSNPDFTKCFONTVLVWVPCFYLMACPEFFLYLSRH 60
QY 61 DRGIIQMTPLNKRTTALGFLIMTYCWDLFPSFWERSRGIFLAVFLVSPLLGITTLLA 120
DB 61 DRGIIQMTPLNKRTTALGFLIMTYCWDLFPSFWERSRGIFLAVFLVSPLLGITTLLA 120
QY 121 TFLIQERRRKQVOSSGIMLTFEVALVICALAIIRSKIMTALKEDAOYDLPRDITFFYYFS 180
DB 121 TFLIQERRRKQVOSSGIMLTFEVALVICALAIIRSKIMTALKEDAOYDLPRDITFFYYFS 180
QY 121 TFLIQERRRKQVOSSGIMLTFEVALVICALAIIRSKIMTALKEDAOYDLPRDITFFYYFS 180
DB 121 TFLIQERRRKQVOSSGIMLTFEVALVICALAIIRSKIMTALKEDAOYDLPRDITFFYYFS 180
QY 181 LLLIQLVLSGFSRSPFLFSETHIDPNCPRESSASFLSRITTFWITGLIVRGYRPLEGSD 240
DB 181 LLLIQLVLSGFSRSPFLFSETHIDPNCPRESSASFLSRITTFWITGLIVRGYRPLEGSD 240
QY 241 LMSLNKEDTSGOAVPVLYKMKKKECATRROPVYVYSSDPAOPKSSKYDAEEVAL 300
DB 241 LMSLNKEDTSGOAVPVLYKMKKKECATRROPVYVYSSDPAOPKSSKYDAEEVAL 300
QY 301 IVKSPQKEMNPDLFKVLYKTGPFYFLMSFFPKAIHDLIMFSGPQILKLLIKFVNDTKAPD 360
DB 301 IVKSPQKEMNPDLFKVLYKTGPFYFLMSFFPKAIHDLIMFSGPQILKLLIKFVNDTKAPD 360
QY 361 WQGFYTVLLFVYACIQTVLHQVFHLCFVSGMKRIKTAIVGAYVRKALVITNSARKSSTV 420
DB 361 WQGFYTVLLFVYACIQTVLHQVFHLCFVSGMKRIKTAIVGAYVRKALVITNSARKSSTV 420

```

```

QY 421 GEIVNLSVDAQRFMDLATYINMISAPLOYIALYLLMLNIGPSVLAVGAVMVLMPVN 480
DB 421 GEIVNLSVDAQRFMDLATYINMISAPLOYIALYLLMLNIGPSVLAVGAVMVLMPVN 480
QY 481 AVAMAKTKTYOVAMHKSNDRIKLMNELNGIKYLKYLAMELAPDKYLAIROELKYLK 540
DB 481 AVAMAKTKTYOVAMHKSNDRIKLMNELNGIKYLKYLAMELAPDKYLAIROELKYLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVVTIDENNIILDAQTAFAVSLFENILRPPLNLP 600
DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVVTIDENNIILDAQTAFAVSLFENILRPPLNLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
QY 661 LINGITFSPREGALVAVVQVCGKSLLSALLAENDKVEGHVALKGSVAVYPOQAMIQND 720
DB 661 LINGITFSPREGALVAVVQVCGKSLLSALLAENDKVEGHVALKGSVAVYPOQAMIQND 720
QY 721 SLRENILFCQLEEPYRSVIOACALPDLLELPDSEGRTEIGEKGVNLSCGQKORVSLAR 780
DB 721 SLRENILFCQLEEPYRSVIOACALPDLLELPDSEGRTEIGEKGVNLSCGQKORVSLAR 780
QY 781 AVSNADITYLFDPLSAVDAHVKHIFENVIGPKMLNKXTRILVTHSMSTLPQVDVILV 840
DB 781 AVSNADITYLFDPLSAVDAHVKHIFENVIGPKMLNKXTRILVTHSMSTLPQVDVILV 840
QY 841 MSGKISMSGSOELLARDGAFAEFLRTYASTEOQDAENGVTGVSQGEAKQEMGM 900
DB 841 MSGKISMSGSOELLARDGAFAEFLRTYASTEOQDAENGVTGVSQGEAKQEMGM 900
QY 901 LVTSAGKOLOROLSSSSSSSGDISRHNSSTAELKKAKEETWKLMEADKAQOGYKL 960
DB 901 LVTSAGKOLOROLSSSSSSSGDISRHNSSTAELKKAKEETWKLMEADKAQOGYKL 960
QY 961 SVYWDYMAKIGLFIISFISIFLPMCNHVASALSNYMLSLMTDDPIVNGOEHKXVLSYVG 1020
DB 961 SVYWDYMAKIGLFIISFISIFLPMCNHVASALSNYMLSLMTDDPIVNGOEHKXVLSYVG 1020
QY 1021 ALGISQGIAPVGYSAVSIIGGILARCLHVDLHLSILSPMSFFERTSSGNLVNFSKEL 1080
DB 1021 ALGISQGIAPVGYSAVSIIGGILARCLHVDLHLSILSPMSFFERTSSGNLVNFSKEL 1080
QY 1081 DTVDSMIEVIMKFMFSGLPFNVGACIVILATPIAIIIPPLGLYFFVQRPYVASSROL 1140
DB 1081 DTVDSMIEVIMKFMFSGLPFNVGACIVILATPIAIIIPPLGLYFFVQRPYVASSROL 1140
QY 1141 KRLESVSRSPYSHFNETLLGVSVIRAFEEORFIHQSDLKVDENOKAYYPSIVANRWLA 1200
DB 1141 KRLESVSRSPYSHFNETLLGVSVIRAFEEORFIHQSDLKVDENOKAYYPSIVANRWLA 1200
QY 1201 VRLEVCVNCIYLFALLFAVISHSLSAGLVGVSYSIQVTTYLWLVYRMSSEMTNIVA 1260
DB 1201 VRLEVCVNCIYLFALLFAVISHSLSAGLVGVSYSIQVTTYLWLVYRMSSEMTNIVA 1260
QY 1261 VERLKEVSETEKEAPMOIOETRPSPSPQVGRVERRNCLARREDLDFVLRHINVTINGG 1320
DB 1261 VERLKEVSETEKEAPMOIOETRPSPSPQVGRVERRNCLARREDLDFVLRHINVTINGG 1320
QY 1321 EKVGVIGRTGAGKSSLLGLFRINESABGEIITIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
DB 1321 EKVGVIGRTGAGKSSLLGLFRINESABGEIITIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
QY 1381 SGSLRPMNDPFSQYSDERWMTSLBLAHLKDFVSALPDKLDHECAGGENLSVGQOLVCL 1440
DB 1381 SGSLRPMNDPFSQYSDERWMTSLBLAHLKDFVSALPDKLDHECAGGENLSVGQOLVCL 1440
QY 1441 ABALIRKTKIIVLDATAAVDLETDLIQSTIRTPBEDCTVLTIAHRLNTIMDYTRVILV 1500
DB 1441 ABALIRKTKIIVLDATAAVDLETDLIQSTIRTPBEDCTVLTIAHRLNTIMDYTRVILV 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMADAGLV 1531

```

```

DB 1501 DKGEIOEYGAPSDLLQORGLFYSMADAGLV 1531
|||||
RESULT 11
US-08-463-179A-2
; Sequence 2, Application US/08463179A
; Patent No. 6001563
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P. C.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,179A
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A. Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: P01-002CP8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-463-179A-2
Query Match 80.6%; Score 7849; DB 2; length 1531;
Best local similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MALGFGSADSDPLMDNNTVWNTSNPDFTKCFONTVYLVWPCFYLVACFPFYLYSRH 60
DB 1 MALGFGSADSDPLMDNNTVWNTSNPDFTKCFONTVYLVWPCFYLVACFPFYLYSRH 60
QY 61 DRGYIOMPLANKTKTALGFLMIYCMADLFYSFMRSGIFLAIVFLVSPILGITTLLA 120
DB 61 DRGYIOMPLANKTKTALGFLMIYCMADLFYSFMRSGIFLAIVFLVSPILGITTLLA 120
QY 121 TFLIQLERRKGVQSGIMLTFMLVALVCALILRSKIMTALKEDAQVDLFPDITFYVYS 180
DB 121 TFLIQLERRKGVQSGIMLTFMLVALVCALILRSKIMTALKEDAQVDLFPDITFYVYS 180
QY 181 LLLIQVLVSCFSDSPFLSETIHDNPNCPRESSASFLSITTWMTGLIVRGYROPLESD 240
DB 181 LLLIQVLVSCFSDSPFLSETIHDNPNCPRESSASFLSITTWMTGLIVRGYROPLESD 240
QY 240 LLLIQVLVSCFSDSPFLSETIHDNPNCPRESSASFLSITTWMTGLIVRGYROPLESD 240
DB 240 LLLIQVLVSCFSDSPFLSETIHDNPNCPRESSASFLSITTWMTGLIVRGYROPLESD 240

```

QY 241 LMSLNKEDTSEQVAVLVKNNKKECAKTRKQPVKVVYSSKDPAQPKESSKYDANEVEAL 300
 DB 241 LMSLNKEDTSEQVAVLVKNNKKECAKTRKQPVKVVYSSKDPAQPKESSKYDANEVEAL 300
 QY 301 IVKSPQKEMNDSLFEVLYKTFGPFYFLMSFFPKAIDHLMFSGPOLIKLIFVNDTKAPD 360
 DB 301 IVKSPQKEMNDSLFEVLYKTFGPFYFLMSFFPKAIDHLMFSGPOLIKLIFVNDTKAPD 360
 QY 361 WQGFYVYVLLFVTACLOTLVHOYFHCVFVSGMRTKTAIVGAVYRKALVTNSARKSSTV 420
 DB 361 WQGFYVYVLLFVTACLOTLVHOYFHCVFVSGMRTKTAIVGAVYRKALVTNSARKSSTV 420
 QY 421 GEIVNLMSVDQRFMDLATYINMTWSAPLOVILALYLLMLNGPSVLGAVVWMLMPVN 480
 DB 421 GEIVNLMSVDQRFMDLATYINMTWSAPLOVILALYLLMLNGPSVLGAVVWMLMPVN 480
 QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVYLKAYMELAFDKVLAIROBELKYLK 540
 DB 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVYLKAYMELAFDKVLAIROBELKYLK 540
 QY 541 KSAVLSAVGTFTWCTPPLVALCTPAVYVYTDENNILDAQTAFLPILRPLNLP 600
 DB 541 KSAVLSAVGTFTWCTPPLVALCTPAVYVYTDENNILDAQTAFLPILRPLNLP 600
 QY 601 MVISSIVQASVSLKRLRIFLSHEELBPSIERRPVKDGGGNSITVRNATFTMARSDPT 660
 DB 601 MVISSIVQASVSLKRLRIFLSHEELBPSIERRPVKDGGGNSITVRNATFTMARSDPT 660
 QY 661 LMGITFSP1PEGALVAVVGQVCGKSSLSALLAEMDKVGHVAKGSAAVYPOQAWIOND 720
 DB 661 LMGITFSP1PEGALVAVVGQVCGKSSLSALLAEMDKVGHVAKGSAAVYPOQAWIOND 720
 QY 721 SLRNNIIFGCOLBEPYRSVYQACALLPDLLEILPSGDRTEIGEKVNLSSGQKRVSLAR 780
 DB 721 SLRNNIIFGCOLBEPYRSVYQACALLPDLLEILPSGDRTEIGEKVNLSSGQKRVSLAR 780
 QY 781 AVVSNADITLFPDDLSAVDAHVGHIFENVIGPKMLKNKRIIVTMSMSTLPQVDYIV 840
 DB 781 AVVSNADITLFPDDLSAVDAHVGHIFENVIGPKMLKNKRIIVTMSMSTLPQVDYIV 840
 QY 841 MSGGKISEMGSYQELLRADGAFAEFLRTYASTEODAEENGYTVSGPGKEAQMENG 900
 DB 841 MSGGKISEMGSYQELLRADGAFAEFLRTYASTEODAEENGYTVSGPGKEAQMENG 900
 QY 901 LVTDSAGKOLQROUSSSSSYSGDISRHHNSTABLQKAEAKKEETWKLMEADKAQTGYKL 960
 DB 901 LVTDSAGKOLQROUSSSSSYSGDISRHHNSTABLQKAEAKKEETWKLMEADKAQTGYKL 960
 QY 961 SYVWDYMKATGLFISFLSTPLFMCNHVSALASNTWLSLMTDDPIVNGTQEHTRKLSVYG 1020
 DB 961 SYVWDYMKATGLFISFLSTPLFMCNHVSALASNTWLSLMTDDPIVNGTQEHTRKLSVYG 1020
 QY 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLHSLIRSPMSFEERTPSGULVNFSEKL 1080
 DB 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLHSLIRSPMSFEERTPSGULVNFSEKL 1080
 QY 1081 DTVDMSIPEVILKMFMSLFNVIGACIVILATPIAIIIPGLIYFFVQRFYVASSROL 1140
 DB 1081 DTVDMSIPEVILKMFMSLFNVIGACIVILATPIAIIIPGLIYFFVQRFYVASSROL 1140
 QY 1141 KRLESVRSRPYSHFNFTLGVSTIRAPBEBERFIHOSDLKVDENQAKYPSIYANWLA 1200
 DB 1141 KRLESVRSRPYSHFNFTLGVSTIRAPBEBERFIHOSDLKVDENQAKYPSIYANWLA 1200
 QY 1201 VRLECVNGCIYLFALFAVISRHSLSAGLVGLSVSYSLQVTTYANMLVRMSSEMETNIVA 1260
 DB 1201 VRLECVNGCIYLFALFAVISRHSLSAGLVGLSVSYSLQVTTYANMLVRMSSEMETNIVA 1260
 QY 1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVEFRNYCLRYREDLDFVLRIHINVTNGG 1320
 DB 1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVEFRNYCLRYREDLDFVLRIHINVTNGG 1320
 QY 1321 EKVGIVERTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLAFKTIITIPQDVL 1380

DB 1321 EKVGIVERTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLAFKTIITIPQDVL 1380
 QY 1381 SGSLRNMULDPSQYSDDEEWTSLBLAHLKOPVSALPKLHCEAGGEMNLSVGOROLVCL 1440
 DB 1381 SGSLRNMULDPSQYSDDEEWTSLBLAHLKOPVSALPKLHCEAGGEMNLSVGOROLVCL 1440
 QY 1441 ARALLRKTILVDEATAVADLETFDDLIQSTIRTFQFEDCTVLTARHLNFTIMDYTRIVL 1500
 DB 1441 ARALLRKTILVDEATAVADLETFDDLIQSTIRTFQFEDCTVLTARHLNFTIMDYTRIVL 1500
 QY 1501 DKGEIOEYGAPSDILQORGLFYSMAXKADAGLV 1531
 DB 1501 DKGEIOEYGAPSDILQORGLFYSMAXKADAGLV 1531
 RESULT 12
 US-08-461-384B-2
 : Sequence 2, Application US/08461384B
 : Patent No. 6025473
 : GENERAL INFORMATION:
 : APPLICANT: Cole, Susan P.C.
 : APPLICANT: Deeley, Roger G.
 : TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
 : NUMBER OF SEQUENCES: 10
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: PARTQ RESEARCH & DEVELOPMENT INNOVATIONS
 : STREET: Queen's University at Kingston
 : CITY: Kingston
 : STATE: Ontario
 : COUNTRY: CANADA
 : ZIP: K7L 3N6
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: ASCII text
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/461,384B
 : FILING DATE: 05-JUN-95
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 07/966,923
 : FILING DATE: 27-OCT-1992
 : APPLICATION NUMBER: 08/029,340
 : FILING DATE: 8-MAR-1993
 : APPLICATION NUMBER: 08/141,893
 : FILING DATE: 26-OCT-1993
 : APPLICATION NUMBER: 08/407,207
 : FILING DATE: 20-MAR-1995
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Steeg, Carol Miernicki
 : REGISTRATION NUMBER: 39,539
 : REFERENCE/DOCKET NUMBER: 01547
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (613) 545-2342
 : TELEFAX: (613) 545-6853
 : INFORMATION FOR SEQ ID NO: 2:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1531 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : US-08-461-384B-2
 Query Match 80.64; Score 7849; DB 2; Length 1531;
 Best Local Similarity 99.94; Pred. No. 0;
 Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MALRGFCSADGSDPLMDMNTWNTSNPDFTKCFQNTVLVWPCFYIMACFPFYFLYSRH 60
 DB 1 MALRGFCSADGSDPLMDMNTWNTSNPDFTKCFQNTVLVWPCFYIMACFPFYFLYSRH 60
 QY 61 DRGIVQMTPLNKTKTALGFLMIVCANADLFYSEWERSRGIFLAPVFLVSTLLGITTLLA 120

```

Db      61 DRGIQWTPLNKKTALGFLLMIVCWLDFYSFWERSGIFLAPFLVSPILLGITLLA 120
Qy      121 TFLQLERRKGVSSGIMLTFMVALVCALAIIRSKITALTALKEDAQVDLPFDITFYVFS 180
Db      121 TFLQLERRKGVSSGIMLTFMVALVCALAIIRSKITALTALKEDAQVDLPFDITFYVFS 180
Qy      181 LLLIQLVLSGSDSPLESETIHDNPNCPRESSASFLSRITFMWITGLIVRGYROPLEGSD 240
Db      181 LLLIQLVLSGSDSPLESETIHDNPNCPRESSASFLSRITFMWITGLIVRGYROPLEGSD 240
Qy      241 LMSLNKEDTSEQVVPVIVKMKKGCATRKQPVKVVYSSKDPAPQKESKTDANEVVAL 300
Db      241 LMSLNKEDTSEQVVPVIVKMKKGCATRKQPVKVVYSSKDPAPQKESKTDANEVVAL 300
Qy      301 IVKSPQKEMNPSLRFVLKTFGFPYFLMSFPFKAIHDLWMFSGPOLIKLILFVNDTKAPD 360
Db      301 IVKSPQKEMNPSLRFVLKTFGFPYFLMSFPFKAIHDLWMFSGPOLIKLILFVNDTKAPD 360
Qy      361 WQGYFYTVLFTVTAQLTLVHLQYFPHICFVSGMRIKTAVIGA VYRKALVITNSARKSSTV 420
Db      361 WQGYFYTVLFTVTAQLTLVHLQYFPHICFVSGMRIKTAVIGA VYRKALVITNSARKSSTV 420
Qy      421 GEIYNLMSVDAORFMDLATYINMTWSAPLOVILAYLMLNGPSVLGAVVMVLMVEVN 480
Db      421 GEIYNLMSVDAORFMDLATYINMTWSAPLOVILAYLMLNGPSVLGAVVMVLMVEVN 480
Qy      481 AVMAKTKTYOVAAHMKSKDNRIKLMNEILNGIKVCLKLAMELAPKDYLAIROBELKYLK 540
Db      481 AVMAKTKTYOVAAHMKSKDNRIKLMNEILNGIKVCLKLAMELAPKDYLAIROBELKYLK 540
Qy      541 KSAVLSAVGTFWCTPFLVALCTFAVYVITDENNILLDAQAFVSLAFNILLRPELNLIP 600
Db      541 KSAVLSAVGTFWCTPFLVALCTFAVYVITDENNILLDAQAFVSLAFNILLRPELNLIP 600
Qy      601 MVISSIVASVSLKRLIFLSHEBELPDSIERRPYKDGCGTNSITVRAATFMARSDDPT 660
Db      601 MVISSIVASVSLKRLIFLSHEBELPDSIERRPYKDGCGTNSITVRAATFMARSDDPT 660
Qy      661 LINGITFSPREGALVAVVGOVGGKSLLSALLAEMDKVEGHAIKGSVAVYPOQAMIOND 720
Db      661 LINGITFSPREGALVAVVGOVGGKSLLSALLAEMDKVEGHAIKGSVAVYPOQAMIOND 720
Qy      721 SLRENILFGCOLBEPYRSVIOACALLPDLBILPSGDRTEIGEKNVLSGGQKORVSLAR 780
Db      721 SLRENILFGCOLBEPYRSVIOACALLPDLBILPSGDRTEIGEKNVLSGGQKORVSLAR 780
Qy      781 AVYSNADIIYLPDDPLSAVDAAHVGKHI FENVIGPKMKNKTRILVTHSMSTYLPQVDVILV 840
Db      781 AVYSNADIIYLPDDPLSAVDAAHVGKHI FENVIGPKMKNKTRILVTHSMSTYLPQVDVILV 840
Qy      841 MSGGKISMGVSOELLARDGAFAEFLRTYASTEOGDAEBENVTVSGPKKAKOMENGM 900
Db      841 MSGGKISMGVSOELLARDGAFAEFLRTYASTEOGDAEBENVTVSGPKKAKOMENGM 900
Qy      901 LVTTDSAGKOLOROLSSSSSSSGDISRHHNSTAELOKAKAKKEETWKLMEADYACTGOYKL 960
Db      901 LVTTDSAGKOLOROLSSSSSSSGDISRHHNSTAELOKAKAKKEETWKLMEADYACTGOYKL 960
Qy      961 SVYMDYMAKIGLFLISFLSIFLPMCNHVSALASNYWLSMTDDPIVNGTQETHKVALSVYG 1020
Db      961 SVYMDYMAKIGLFLISFLSIFLPMCNHVSALASNYWLSMTDDPIVNGTQETHKVALSVYG 1020
Qy      1021 ALGISQGIAVFGYMAVSIIGLILASRCLHVDLHLSILSPMSFPERTSGNLYNFSKEL 1080
Db      1021 ALGISQGIAVFGYMAVSIIGLILASRCLHVDLHLSILSPMSFPERTSGNLYNFSKEL 1080
Qy      1081 DTVDMSIEVIMKFMGSLFNVIGACIVILLATPIAIIIPPLGILYFFQRFYVASSRQL 1140
Db      1081 DTVDMSIEVIMKFMGSLFNVIGACIVILLATPIAIIIPPLGILYFFQRFYVASSRQL 1140
Qy      1141 KRLBSVSSPYVSHNETLLGVSVIRAEEOERFIHQSDLKVDENOKAYPSIVANRWLA 1200

```

```

Db      1141 KRLBSVSSPYVSHNETLLGVSVIRAEEOERFIHQSDLKVDENOKAYPSIVANRWLA 1200
Qy      1201 VRLCEVGNCTIVLFAALFAVISRHSLSAGLVGLSVYSLOVTTYLNMVLRMSSEMETNIVA 1260
Db      1201 VRLCEVGNCTIVLFAALFAVISRHSLSAGLVGLSVYSLOVTTYLNMVLRMSSEMETNIVA 1260
Qy      1261 VERLKEYSETEKEAPWMOIOETAPSSWPQVGRVFNRYCIRYRDLDPVLRHINVTINGG 1320
Db      1261 VERLKEYSETEKEAPWMOIOETAPSSWPQVGRVFNRYCIRYRDLDPVLRHINVTINGG 1320
Qy      1321 EKVGIVGRTAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
Db      1321 EKVGIVGRTAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
Qy      1381 SGLSRMNLDPFSQYSDSEVWTSLELAHLKDFVSAIPDKLDHECAGEGENTSVGROLVCL 1440
Db      1381 SGLSRMNLDPFSQYSDSEVWTSLELAHLKDFVSAIPDKLDHECAGEGENTSVGROLVCL 1440
Qy      1441 ARALLRTKTLIVDEATAVADLEFTDILLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRYVL 1500
Db      1441 ARALLRTKTLIVDEATAVADLEFTDILLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRYVL 1500
Qy      1501 DKGEIOEYGAPSDLLQORGLFYSMAXDAGLV 1531
Db      1501 DKGEIOEYGAPSDLLQORGLFYSMAXDAGLV 1531

```

RESULT 13

US-08-407-207A-2

Sequence 2, Application US/08407207A

Patent No. 6063621

GENERAL INFORMATION:

APPLICANT: Deeley, Roger G.

APPLICANT: Cole, Susan P. C.

TITLE OF INVENTION: ANTIBODIES TO A MULTIDRUG RESISTANCE PROTEIN

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESSES:

ADDRESS: PARTO RESEARCH & DEVELOPMENT INNOVATIONS

STREET: Queen's University at Kingston

CITY: Kingston

STATE: Ontario

COUNTRY: CANADA

ZIP: K7L 3N6

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/407,207A

FILING DATE: 20-MAR-1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/966,923

FILING DATE: 27-OCT-1992

APPLICATION NUMBER: 08/029,340

FILING DATE: 8-MAR-1993

APPLICATION NUMBER: 08/141,893

FILING DATE: 26-OCT-1993

ATTORNEY/AGENT INFORMATION:

NAME: Steeg, Carol Miernicki

REGISTRATION NUMBER: 39,539

REFERENCE/DOCKET NUMBER: 01512

TELECOMMUNICATION INFORMATION:

TELEPHONE: (613) 545-2342

TELEFAX: (613) 545-6853

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1531 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-407-207A-2

Query Match 80.6% Score 7849; DB 2; Length 1531;
 Best Local Similarity 99.9% Pred. No. 0;
 Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 MALRGFCGADSDPLMDNVTMTSNPDTKCFONTLVWVPCFYLMACPFYFLYSRHH 60
 1 MALRGFCGADSDPLMDNVTMTSNPDTKCFONTLVWVPCFYLMACPFYFLYSRHH 60
 61 DRGTYQMTPLNKTALGFLMIYCMADLFSPMERSGIFLAPFVLSPILLGITLLA 120
 61 DRGTYQMTPLNKTALGFLMIYCMADLFSPMERSGIFLAPFVLSPILLGITLLA 120
 61 DRGTYQMTPLNKTALGFLMIYCMADLFSPMERSGIFLAPFVLSPILLGITLLA 120
 121 TELLOERKGVSSGIMLTFMILVALYCALAILSKITALKEDAOVDLPFDITFYVFS 180
 121 TELLOERKGVSSGIMLTFMILVALYCALAILSKITALKEDAOVDLPFDITFYVFS 180
 121 TELLOERKGVSSGIMLTFMILVALYCALAILSKITALKEDAOVDLPFDITFYVFS 180
 181 LLLIQLVLSGSDSPLESETIHDNPPCESSASFLSRITFMWITGLIVRGYRPLESSD 240
 181 LLLIQLVLSGSDSPLESETIHDNPPCESSASFLSRITFMWITGLIVRGYRPLESSD 240
 181 LLLIQLVLSGSDSPLESETIHDNPPCESSASFLSRITFMWITGLIVRGYRPLESSD 240
 241 LMSLNKEDTSQVVPVLVKNKKKCAKTRKOPVXVYSSKDPAPKSSKYDANEYVAL 300
 241 LMSLNKEDTSQVVPVLVKNKKKCAKTRKOPVXVYSSKDPAPKSSKYDANEYVAL 300
 241 LMSLNKEDTSQVVPVLVKNKKKCAKTRKOPVXVYSSKDPAPKSSKYDANEYVAL 300
 301 IVKSPKEMNPSLFRVLYKTEGPFYFMSFFKAIHDMFSGPOLIKLLIFVNDTKAPD 360
 301 IVKSPKEMNPSLFRVLYKTEGPFYFMSFFKAIHDMFSGPOLIKLLIFVNDTKAPD 360
 301 IVKSPKEMNPSLFRVLYKTEGPFYFMSFFKAIHDMFSGPOLIKLLIFVNDTKAPD 360
 361 WQGYFYTVLFTVACLOTLVHQYFHLICFVSGMIRKTAIVAGVYRKALVITNSARKSTV 420
 361 WQGYFYTVLFTVACLOTLVHQYFHLICFVSGMIRKTAIVAGVYRKALVITNSARKSTV 420
 361 WQGYFYTVLFTVACLOTLVHQYFHLICFVSGMIRKTAIVAGVYRKALVITNSARKSTV 420
 421 GEIYNLMSVDAQRFMDLATYINMTWSAPLQVILALYILMLNGSVLAGVAVMILAVPN 480
 421 GEIYNLMSVDAQRFMDLATYINMTWSAPLQVILALYILMLNGSVLAGVAVMILAVPN 480
 421 GEIYNLMSVDAQRFMDLATYINMTWSAPLQVILALYILMLNGSVLAGVAVMILAVPN 480
 481 AVMMKTKTYOVAMHMSKDNRIKLMNELINGIKYLKAYAMELARKDKYLAIRBELKYLK 540
 481 AVMMKTKTYOVAMHMSKDNRIKLMNELINGIKYLKAYAMELARKDKYLAIRBELKYLK 540
 481 AVMMKTKTYOVAMHMSKDNRIKLMNELINGIKYLKAYAMELARKDKYLAIRBELKYLK 540
 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNIIDAOAFYSIALFNILRPLNLP 600
 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNIIDAOAFYSIALFNILRPLNLP 600
 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNIIDAOAFYSIALFNILRPLNLP 600
 601 MVISSIVQASVSLKRLRIFLSHEELBPDSIERRPVKOGGJNSITVRNATFTMARSDPT 660
 601 MVISSIVQASVSLKRLRIFLSHEELBPDSIERRPVKOGGJNSITVRNATFTMARSDPT 660
 601 MVISSIVQASVSLKRLRIFLSHEELBPDSIERRPVKOGGJNSITVRNATFTMARSDPT 660
 661 LINGITFSPGALVAVVQVCGCGKSLLSALLAEMDKVGHVAKGSVAVYPOQAMIOND 720
 661 LINGITFSPGALVAVVQVCGCGKSLLSALLAEMDKVGHVAKGSVAVYPOQAMIOND 720
 661 LINGITFSPGALVAVVQVCGCGKSLLSALLAEMDKVGHVAKGSVAVYPOQAMIOND 720
 721 SLRENILFGCOLPEEYRSVIOACALLPDLIELPSGDRTEIGEKNVLSGGOKORVSLAR 780
 721 SLRENILFGCOLPEEYRSVIOACALLPDLIELPSGDRTEIGEKNVLSGGOKORVSLAR 780
 721 SLRENILFGCOLPEEYRSVIOACALLPDLIELPSGDRTEIGEKNVLSGGOKORVSLAR 780
 781 AVYSNADIYLFDDPLSAVDAHVGKHIENVIQPKMLKNKTRILVTHSMSTLPQVDYIIV 840
 781 AVYSNADIYLFDDPLSAVDAHVGKHIENVIQPKMLKNKTRILVTHSMSTLPQVDYIIV 840
 781 AVYSNADIYLFDDPLSAVDAHVGKHIENVIQPKMLKNKTRILVTHSMSTLPQVDYIIV 840
 841 MSGGKISGMSYQELLARDAFAFPLRTYASTBOBDAEENGVTGVSPPGEAKOMENGM 900
 841 MSGGKISGMSYQELLARDAFAFPLRTYASTBOBDAEENGVTGVSPPGEAKOMENGM 900
 841 MSGGKISGMSYQELLARDAFAFPLRTYASTBOBDAEENGVTGVSPPGEAKOMENGM 900
 901 LVYDSAGKOLQROLSSSSSYSGDISRHHNSTAEIQAFAKKEETWKLMEADKAQTQVYL 960
 901 LVYDSAGKOLQROLSSSSSYSGDISRHHNSTAEIQAFAKKEETWKLMEADKAQTQVYL 960
 901 LVYDSAGKOLQROLSSSSSYSGDISRHHNSTAEIQAFAKKEETWKLMEADKAQTQVYL 960
 961 SVYWDYMKAIQGLFISFLSIFLPMGNHVSALASNTWLSIWDOPVINGTOEHTKYRLSVYG 1020
 961 SVYWDYMKAIQGLFISFLSIFLPMGNHVSALASNTWLSIWDOPVINGTOEHTKYRLSVYG 1020
 961 SVYWDYMKAIQGLFISFLSIFLPMGNHVSALASNTWLSIWDOPVINGTOEHTKYRLSVYG 1020

1021 ALGISOGIAYRGVMAVSGIGILLASRCLHVDLHSLIRSPMSFFERTPSGNLVNRSKEL 1080
 1021 ALGISOGIAYRGVMAVSGIGILLASRCLHVDLHSLIRSPMSFFERTPSGNLVNRSKEL 1080
 1081 DTVDMSIPEVIMKMGSLFNVIAGACIVILLATPPIAIIIPPLGIYFFVORFYVAASRQL 1140
 1081 DTVDMSIPEVIMKMGSLFNVIAGACIVILLATPPIAIIIPPLGIYFFVORFYVAASRQL 1140
 1081 DTVDMSIPEVIMKMGSLFNVIAGACIVILLATPPIAIIIPPLGIYFFVORFYVAASRQL 1140
 1141 KRLSVSRSPYSHFNFTLGVSVIRAFEBQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
 1141 KRLSVSRSPYSHFNFTLGVSVIRAFEBQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
 1141 KRLSVSRSPYSHFNFTLGVSVIRAFEBQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
 1201 VRLCEVGNCLVFPALFAVIRSRHLSAGLVGLSVSYSLQVTTYINMLVRRMSSEMETNIVA 1260
 1201 VRLCEVGNCLVFPALFAVIRSRHLSAGLVGLSVSYSLQVTTYINMLVRRMSSEMETNIVA 1260
 1201 VRLCEVGNCLVFPALFAVIRSRHLSAGLVGLSVSYSLQVTTYINMLVRRMSSEMETNIVA 1260
 1261 VERLKEYSETEKEAPMOIOETAPPSMPQVGRVFRNYCLARYEDLDFVLRHINVTINGG 1320
 1261 VERLKEYSETEKEAPMOIOETAPPSMPQVGRVFRNYCLARYEDLDFVLRHINVTINGG 1320
 1261 VERLKEYSETEKEAPMOIOETAPPSMPQVGRVFRNYCLARYEDLDFVLRHINVTINGG 1320
 1321 EKVGIVRTGAGKSLTLGLFRINESABGEIINDGINIAKIGLHDLRFKTIIPQDPVLF 1380
 1321 EKVGIVRTGAGKSLTLGLFRINESABGEIINDGINIAKIGLHDLRFKTIIPQDPVLF 1380
 1321 EKVGIVRTGAGKSLTLGLFRINESABGEIINDGINIAKIGLHDLRFKTIIPQDPVLF 1380
 1381 SGSLRMLDPPSOYSDSEWMTSLAHLKOPVSALPKLHCEAGGSGNLSVGOROLVCL 1440
 1381 SGSLRMLDPPSOYSDSEWMTSLAHLKOPVSALPKLHCEAGGSGNLSVGOROLVCL 1440
 1381 SGSLRMLDPPSOYSDSEWMTSLAHLKOPVSALPKLHCEAGGSGNLSVGOROLVCL 1440
 1441 ABALLRKTILVLDENATAVADLETDDLIQSTIRTOFEDCTVLTAAHRLNTIMDTRVIVL 1500
 1441 ABALLRKTILVLDENATAVADLETDDLIQSTIRTOFEDCTVLTAAHRLNTIMDTRVIVL 1500
 1441 ABALLRKTILVLDENATAVADLETDDLIQSTIRTOFEDCTVLTAAHRLNTIMDTRVIVL 1500
 1501 DKGEOYGAAPSDLLQORGLFYMAKDAGLV 1531
 1501 DKGEOYGAAPSDLLQORGLFYMAKDAGLV 1531
 1501 DKGEOYGAAPSDLLQORGLFYMAKDAGLV 1531

RESULT 14
 US-08-463-092B-6
 Sequence 6, Application US/08463092B
 Patent No. 5766880
 GENERAL INFORMATION:
 APPLICANT: Cole, Susan P. C.
 APPLICANT: Deeley, Roger G.
 TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
 NUMBER OF SEQUENCES: 9
 TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: PARTEO RESEARCH & DEVELOPMENT INNOVATIONS
 STREET: Queen's University at Kingston
 CITY: Kingston
 STATE: Ontario
 COUNTRY: CANADA
 ZIP: K7L 3N6
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/463,092B
 FILING DATE: 05-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/966,923
 FILING DATE: 27-OCT-1992
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/029,340
 FILING DATE: 8-MAR-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/141,893

FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Mernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-092B-6

Query Match 71.9%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRGFCAGSDGSDPLMDNNTWNTSNPDFTKCFONTVLWVWPCFYLMACFPFFLYLSRH 60
DB 1 MALRSFCAGSDGSDPLMDNNTWNTSNPDFTKCFONTVLWVWPCFYLMACFPFFLYLSRH 60
QY 61 DRGTYQMTPLNKTALGFLMTIVCWAFLYFWMRSKGIPLAPVLYSPITLITTLA 120
DB 61 DRGTYQMTPLNKTALGFLMTIVCWAFLYFWMRSKGIPLAPVLYSPITLITTLA 120
QY 121 TFLDLERRKGVQSSGIMLTFWLVALVCALATLRSKIMTALKEAONVDFRDTFYVES 180
DB 121 TFLDLERRKGVQSSGIMLTFWLVALVCALATLRSKIMTALKEAONVDFRDTFYVES 180
QY 121 TFLDLERRKGVQSSGIMLTFWLVALVCALATLRSKIMTALKEAONVDFRDTFYVES 180
DB 121 TFLDLERRKGVQSSGIMLTFWLVALVCALATLRSKIMTALKEAONVDFRDTFYVES 180
QY 181 LLLIQVLVSCPSDRSPFSETIHDNPPCBSSASFLSRTTFWMTGLIVRGYROPLESD 240
DB 181 LVLVQLVLSVSCPSDRSPFSETIHDNPPCBSSASFLSRTTFWMTGLIVRGYROPLESD 240
QY 241 LMSLNKERTSEOVVLYVKNMKECAKTRKQPKVYVYS-SKDPAPKSSSKVDANEVEYA 299
DB 241 LMSLNKERTSEOVVLYVKNMKECAKTRKQPKVYVYS-SKDPAPKSSSKVDANEVEYA 299
QY 300 LMSLNKERTSEOVVLYVKNMKECAKTRKQPKVYVYS-SKDPAPKSSSKVDANEVEYA 300
DB 300 LMSLNKERTSEOVVLYVKNMKECAKTRKQPKVYVYS-SKDPAPKSSSKVDANEVEYA 300
QY 301 LIVSPHOKEMNSLKFVLYKTRGPYFLMSFPFKAIHDMFSGPOLIKLIFVNDTYAP 359
DB 301 LIVSPHOKEMNSLKFVLYKTRGPYFLMSFPFKAIHDMFSGPOLIKLIFVNDTYAP 359
QY 360 DMQGFYTVLLFVTACLOTLVLAHQYFHLCEVSGMRIKTAVIGAVYRKALVITNSARKSST 419
DB 360 DMQGFYTVLLFVTACLOTLVLAHQYFHLCEVSGMRIKTAVIGAVYRKALVITNSARKSST 419
QY 420 VGEIYNLMSVDAQRPMDLATYINMTWSAPLOVITALLYLMLNLSGSVLAAGVAVMYLMPV 479
DB 420 VGEIYNLMSVDAQRPMDLATYINMTWSAPLOVITALLYLMLNLSGSVLAAGVAVMYLMPV 479
QY 480 NAVMAMKTKTYOVAMHMSKDNRIKLMNEIINGIKVLYAMBLAKFDKVLAROBELKVL 539
DB 480 NAVMAMKTKTYOVAMHMSKDNRIKLMNEIINGIKVLYAMBLAKFDKVLAROBELKVL 539
QY 540 KKSAYLSAVGTFTWCTPPLVALCTPAVYVITIDENNILDAQTAFAVSLAFNIRPPLNL 599
DB 540 KKSAYLSAVGTFTWCTPPLVALCTPAVYVITIDENNILDAQTAFAVSLAFNIRPPLNL 599
QY 541 KKSAYLSAVGTFTWCTPPLVALCTPAVYVITIDENNILDAQTAFAVSLAFNIRPPLNL 600
DB 541 KKSAYLSAVGTFTWCTPPLVALCTPAVYVITIDENNILDAQTAFAVSLAFNIRPPLNL 600
QY 600 PMVYSISYQASVSLKRLIFLSHEBELPDSIRRPYKCGGNSITVRNATTTMARSDPP 659
DB 600 PMVYSISYQASVSLKRLIFLSHEBELPDSIRRPYKCGGNSITVRNATTTMARSDPP 659
QY 601 PMVYSISYQASVSLKRLIFLSHEBELPDSIRRPYKCGGNSITVRNATTTMARSDPP 659
DB 601 PMVYSISYQASVSLKRLIFLSHEBELPDSIRRPYKCGGNSITVRNATTTMARSDPP 659
QY 660 TLNGITFSPREGALVAVVGCGKSSLLSALLAENDKVEGHVTLKGSVAVYPOQAMION 719
DB 660 TLNGITFSPREGALVAVVGCGKSSLLSALLAENDKVEGHVTLKGSVAVYPOQAMION 719

QY 720 DSLRENTLFGCQLEPPYRYVYQACALLPDLEILLPSGDRTEIGRKVNLSSGQKQRYSLA 779
DB 720 DSLRENTLFGHPOENNYKAKMAECALLPDLEILLPSGDRTEIGRKVNLSSGQKQRYSLA 779
QY 780 RAVYSNADITYLFDPLSAVDHVGKHLFENVIGPKMLKNKTRILYTHSMYSLPOVDYI 839
DB 780 RAVYSNADITYLFDPLSAVDHVGKHLFENVIGPKMLKNKTRILYTHSMYSLPOVDYI 839
QY 840 VMSGKXISEMGSYOELLARDGAFAPLRTYASTEOBDAENGYGVSGPKCKAKOMENG 899
DB 840 VMSGKXISEMGSYOELLARDGAFAPLRTYASTEOBDAENGYGVSGPKCKAKOMENG 899
QY 900 MLVTDASGKOLQOLSSSSSYSGDISRHNSSTAELQALBAKCBETWKLMEADKQOTGVK 959
DB 900 MLVTDASGKOLQOLSSSSSYSGDISRHNSSTAELQALBAKCBETWKLMEADKQOTGVK 959
QY 960 LSVYWDYWKALGLISPLSIFLFCNHNVSALASNYMLSLMTDD-PIYNGOEHKRVLSY 1018
DB 960 LSVYWDYWKALGLISPLSIFLFCNHNVSALASNYMLSLMTDD-PIYNGOEHKRVLSY 1018
QY 956 LSVYWNWKALGLITPLSIFLFCNHNVSALASNYMLSLMTDD-PIYNGOEHKRVLSY 1015
DB 956 LSVYWNWKALGLITPLSIFLFCNHNVSALASNYMLSLMTDD-PIYNGOEHKRVLSY 1015
QY 1019 YGALGISQGLAVFGYSMAVSIIGIILASRCLHVDLHSLRSNMSFPFERTPSGNTVNRPSK 1078
DB 1019 YGALGISQGLAVFGYSMAVSIIGIILASRCLHVDLHSLRSNMSFPFERTPSGNTVNRPSK 1078
QY 1016 YGALGIIQGAALIFGYMAVSIIGIIFASRHLHDLVNLVLRSPMSFPFERTPSGNTVNRPSK 1075
DB 1016 YGALGIIQGAALIFGYMAVSIIGIIFASRHLHDLVNLVLRSPMSFPFERTPSGNTVNRPSK 1075
QY 1079 ELDTVDSMTPEVITPMFNGSLFNVIGACTVILATPILAIIPDGLIYFFQRYVYASR 1138
DB 1079 ELDTVDSMTPEVITPMFNGSLFNVIGACTVILATPILAIIPDGLIYFFQRYVYASR 1138
QY 1139 QLKRLSESVRSPPVSHNETLLGVSIRAPEEORFPHOSDLKVDENOKAYPSIVANRW 1198
DB 1139 QLKRLSESVRSPPVSHNETLLGVSIRAPEEORFPHOSDLKVDENOKAYPSIVANRW 1198
QY 1136 QLKRLSESVRSPPVSHNETLLGVSIRAPEEORFPHOSDLKVDENOKAYPSIVANRW 1195
DB 1136 QLKRLSESVRSPPVSHNETLLGVSIRAPEEORFPHOSDLKVDENOKAYPSIVANRW 1195
QY 1199 LAVRELCVNCIVFPAALFAVISRHSLSAGLVGSYSLOVTTYLNLVMSSEMETNI 1258
DB 1199 LAVRELCVNCIVFPAALFAVISRHSLSAGLVGSYSLOVTTYLNLVMSSEMETNI 1258
QY 1196 LAVRELCVNCIVFPAALFAVISRHSLSAGLVGSYSLOVTTYLNLVMSSEMETNI 1255
DB 1196 LAVRELCVNCIVFPAALFAVISRHSLSAGLVGSYSLOVTTYLNLVMSSEMETNI 1255
QY 1259 VAVRELCVNCIVFPAALFAVISRHSLSAGLVGSYSLOVTTYLNLVMSSEMETNI 1318
DB 1259 VAVRELCVNCIVFPAALFAVISRHSLSAGLVGSYSLOVTTYLNLVMSSEMETNI 1318
QY 1256 VAVRELCVNCIVFPAALFAVISRHSLSAGLVGSYSLOVTTYLNLVMSSEMETNI 1315
DB 1256 VAVRELCVNCIVFPAALFAVISRHSLSAGLVGSYSLOVTTYLNLVMSSEMETNI 1315
QY 1319 GGEKVGIVGRGTACKSSLTGLFRINSAGEEIIIDGNINAKIGLHDFKITTIPDPV 1378
DB 1319 GGEKVGIVGRGTACKSSLTGLFRINSAGEEIIIDGNINAKIGLHDFKITTIPDPV 1378
QY 1316 GGEKVGIVGRGTACKSSLTGLFRINSAGEEIIIDGNINAKIGLHDFKITTIPDPV 1375
DB 1316 GGEKVGIVGRGTACKSSLTGLFRINSAGEEIIIDGNINAKIGLHDFKITTIPDPV 1375
QY 1379 LFSGSLRNKNDPFGQYSDSEEWTSLELAHLKD FVSALPDKLDHCAEGENLSVGORQV 1438
DB 1379 LFSGSLRNKNDPFGQYSDSEEWTSLELAHLKD FVSALPDKLDHCAEGENLSVGORQV 1438
QY 1376 LFSGSLRNKNDPFGQYSDSEEWTSLELAHLKD FVSALPDKLDHCAEGENLSVGORQV 1435
DB 1376 LFSGSLRNKNDPFGQYSDSEEWTSLELAHLKD FVSALPDKLDHCAEGENLSVGORQV 1435
QY 1439 CLARALRKTYILVDEATAVADLETDDLIOSTRTOFEDCTVLTIAHRLNTINDYRVI 1498
DB 1439 CLARALRKTYILVDEATAVADLETDDLIOSTRTOFEDCTVLTIAHRLNTINDYRVI 1498
QY 1436 CLARALRKTYILVDEATAVADLETDDLIOSTRTOFEDCTVLTIAHRLNTINDYRVI 1495
DB 1436 CLARALRKTYILVDEATAVADLETDDLIOSTRTOFEDCTVLTIAHRLNTINDYRVI 1495
QY 1499 VLDKGEIOEGAPBDLLOQRGLFYSMAKDALV 1531
DB 1499 VLDKGEIOEGAPBDLLOQRGLFYSMAKDALV 1531
QY 1496 VLDKGEIOEGAPBDLLOQRGLFYSMAKDALV 1528
DB 1496 VLDKGEIOEGAPBDLLOQRGLFYSMAKDALV 1528

RESULT 15
US-08-462-109A-6
Sequence 6, Application US/08462109A
Patent No. 5882875
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Dealey, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA

```
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,109A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: PQI-002CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-6

Query Match      71.9%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRFGSADGSDPLDMNNTVNTSNDPTKCFONTVLVWPCFYLMACEPFYLYSRH 60
DB 1 MALRFSFCSADGSDPLDMNNTVNTSNDPTKCFONTVLVWPCFYLMACEPFYLYSRH 60
QY 61 DRGIIOMTLPNKTKTALGFLMTIYCWADLFYSFWERSRGITFLAVPLVSPILGITTLLA 120
DB 61 DRGIIOMTLPNKTKTALGFLMTIYCWADLFYSFWERSRGITFLAVPLVSPILGITTLLA 120
QY 121 TFLIOLERRKGVSSGIMLTFMLVALYCALAILRSKIMTALKEPAOYDLFRDITFYVFS 180
DB 121 TFLIOLERRKGVSSGIMLTFMLVALYCALAILRSKIMTALKEPAOYDLFRDITFYVFS 180
QY 181 LLLIQLVLSGCSDRSPLFSETIHDNPPCESSASFLSRITFMWITGLIVRGYRQPLEGSD 240
DB 181 LLLIQLVLSGCSDRSPLFSETIHDNPPCESSASFLSRITFMWITGLIVRGYRQPLEGSD 240
QY 241 LMSLNKEDTSEOVVPLVKNMKKECAKTRQKQPVVVS-SKDPAPKESSVVDNEVEEA 299
DB 241 LMSLNKEDTSEOVVPLVKNMKKECAKTRQKQPVVVS-SKDPAPKESSVVDNEVEEA 299
QY 300 LIVESPOKEMNPISLFKLYKTGPGYFLMSFFKAIHDLMPGSGOILKLIKFNVDTKAP 359
DB 300 LIVESPOKEMNPISLFKLYKTGPGYFLMSFFKAIHDLMPGSGOILKLIKFNVDTKAP 359
QY 360 DWQGYFTVLLFTVACIQTLLVLAHQYFHI CTVSGMRITKTAIVIGAVYRRALVITNSARKST 419
DB 360 DWQGYFTVLLFTVACIQTLLVLAHQYFHI CTVSGMRITKTAIVIGAVYRRALVITNSARKST 419
QY 420 VGEIVNLMSYDAQGFMDLATYINNIWISAPLOVITLALVLMNLGSPSLAGAVVWVLMVPV 479
DB 420 VGEIVNLMSYDAQGFMDLATYINNIWISAPLOVITLALVLMNLGSPSLAGAVVWVLMVPV 479
QY 480 NAYVAMKTKTYQVAHMSKDNRIKLMEIINGIKIVLKLXYAMELAFKQVLAIRQEBELKVL 539
DB 480 NAYVAMKTKTYQVAHMSKDNRIKLMEIINGIKIVLKLXYAMELAFKQVLAIRQEBELKVL 539
QY 481 NAYVAMKTKTYQVAHMSKDNRIKLMEIINGIKIVLKLXYAMELAFKQVLAIRQEBELKVL 540
DB 481 NAYVAMKTKTYQVAHMSKDNRIKLMEIINGIKIVLKLXYAMELAFKQVLAIRQEBELKVL 540
```

```
QY 540 KKSAYLSANGFTFWCTPELVALCTFAVYVYTIIDENNILDAQTAFAVSLAFNIIAREPILNL 599
DB 541 KKSAYLSANGFTFWCTPELVALCTFAVYVYTIIDENNILDAQTAFAVSLAFNIIAREPILNL 600
QY 600 PMVSSIVQASVSLKRIKRIPLSHBELPPDSTERBPVDDGGTNSITVRNATFTVARSDDP 659
DB 601 PMVSSIVQASVSLKRIKRIPLSHBELPPDSTERBPVDDGGTNSITVRNATFTVARSDDP 659
QY 660 TLNGITSIPEGALVAVVGQVCGKSSILSLALMEKQVEGHVAIKGSVAVPOQAMION 719
DB 660 TLNGITSIPEGALVAVVGQVCGKSSILSLALMEKQVEGHVAIKGSVAVPOQAMION 719
QY 720 DLSLRNLLFGCOLLEPPYRSVIOACALLPDLIILPSPDRTRIGKGVNLSSGQKQVSLA 779
DB 720 DLSLRNLLFGCOLLEPPYRSVIOACALLPDLIILPSPDRTRIGKGVNLSSGQKQVSLA 779
QY 780 RAYVSNADITLFDPLSAVDPAHVGHKIFENVIGPKMLKNTKRLILVTHSMGYLPQVDYII 839
DB 780 RAYVSNADITLFDPLSAVDPAHVGHKIFENVIGPKMLKNTKRLILVTHSMGYLPQVDYII 839
QY 840 VMSGKISMGSYGELLARDGAPAFPLRTYASTEQEOBDAEENGTVGVSGPKKAKOMENG 899
DB 840 VMSGKISMGSYGELLARDGAPAFPLRTYASTEQEOBDAEENGTVGVSGPKKAKOMENG 899
QY 900 MLVTDGAKQOLQOLSSSSSSGDI SHHNSSTAELQAEAKKETWKLMEADKQOTGVK 959
DB 900 MLVTDGAKQOLQOLSSSSSSGDI SHHNSSTAELQAEAKKETWKLMEADKQOTGVK 959
QY 960 LSVYWDYMKAIIGLIFSLFSLIFLFCNVHVSALASVYMLSLMTDD-PINGTOEHKIVLSV 1018
DB 960 LSVYWDYMKAIIGLIFSLFSLIFLFCNVHVSALASVYMLSLMTDD-PINGTOEHKIVLSV 1018
QY 996 LSVYWDYMKAIIGLIFSLFSLIFLFCNVHVSALASVYMLSLMTDD-PINGTOEHKIVLSV 1018
DB 996 LSVYWDYMKAIIGLIFSLFSLIFLFCNVHVSALASVYMLSLMTDD-PINGTOEHKIVLSV 1018
QY 1019 YGALGISQGIIVGSGYMAVSIIGLILSRCHVDLHLSILKSPMSFFERTSGNLYNRFSK 1078
DB 1016 YGALGISQGIIVGSGYMAVSIIGLILSRCHVDLHLSILKSPMSFFERTSGNLYNRFSK 1075
QY 1079 ELDTVDSMIDPEVIMKMGSLPNVYIGACIYLLATPIAIIIPGLIYFFVQRFYVASSR 1138
DB 1076 ELDTVDSMIDPEVIMKMGSLPNVYIGACIYLLATPIAIIIPGLIYFFVQRFYVASSR 1135
QY 1139 QLRLESVSRSPPVYSHNETLLGVSVIRAFEEQERFTHQSDKVDENQKAYYPSIVANRW 1198
DB 1136 QLRLESVSRSPPVYSHNETLLGVSVIRAFEEQERFTHQSDKVDENQKAYYPSIVANRW 1195
QY 1199 LAVALBECVGCIVLFAALFAVISRHSLSAGLVGSYSIQVTTYLVNMLVRMSSEMETNI 1258
DB 1196 LAVALBECVGCIVLFAALFAVISRHSLSAGLVGSYSIQVTTYLVNMLVRMSSEMETNI 1255
QY 1259 VAVERLKEYSETEKEAPWQIOETAPPSGWPQVGRVEFRNVCARREDLDVFLRHINVTIN 1318
DB 1256 VAVERLKEYSETEKEAPWQIOETAPPSGWPQVGRVEFRNVCARREDLDVFLRHINVTIE 1315
QY 1319 GGEKVGIVGRGTGAKSSTLTGLFRINESAGEIITIGGINIAKIGHDLRKRITIIIPDDPV 1378
DB 1316 GGEKVGIVGRGTGAKSSTLTGLFRINESAGEIITIGGINIAKIGHDLRKRITIIIPDDPV 1375
QY 1379 LPSGSLRMLNDLPPSOYSDSEEVWTSLELAHLKDFVSALPDKLDBHCAEGENLSYGQROLV 1438
DB 1376 LPSGSLRMLNDLPPSOYSDSEEVWTSLELAHLKDFVSALPDKLDBHCAEGENLSYGQROLV 1435
QY 1439 CLAPALIRKTKIIVLDEATVAVDLETDLLIOSTIRFQEDCTVLTIAHRLNTIMDYTRVI 1498
DB 1436 CLAPALIRKTKIIVLDEATVAVDLETDLLIOSTIRFQEDCTVLTIAHRLNTIMDYTRVI 1495
QY 1499 VLDKGEIOEYGAAPSDDLQGRGLFYMAKDKGLV 1531
DB 1496 VLDKGEIOEYGAAPSDDLQGRGLFYMAKDKGLV 1528
```

Search completed: December 15, 2005, 15:49:42
Job time : 48.351 secs

This blank page

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 15:05:00 ; Search time 156.846 Seconds
(without alignments)
5037.528 Million cell updates/sec

Title: US-10-665-283-4

Perfect score: 9734
Sequence: 1 MALRGFCADGSDPLMDMNV.....NTIKVPTPLCTARQLDSDRS 1891

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9734	100.0	1891	US-10-665-283-4	Sequence 4, Appli
2	9734	100.0	1927	US-10-665-283-1	Sequence 1, Appli
3	9734	100.0	1927	US-10-665-283-6	Sequence 6, Appli
4	9704	99.7	1947	US-10-665-283-8	Sequence 8, Appli
5	7860	80.7	1531	US-09-939-853A-86	Sequence 86, Appli
6	7860	80.7	1531	US-10-667-891-6	Sequence 6, Appli
7	7860	80.7	1531	US-10-807-466-6	Sequence 6, Appli
8	7860	80.7	1531	US-10-484-577-678	Sequence 678, App
9	7860	80.7	1531	US-10-889-503-19	Sequence 19, Appli
10	7860	80.7	1531	US-10-756-149-5033	Sequence 5033, Ap
11	7769	79.8	1515	US-09-939-853A-87	Sequence 87, Appli
12	7769	79.8	1515	US-10-618-281-42	Sequence 42, Appli
13	7037	72.3	1388	US-10-408-765A-1718	Sequence 1718, Ap
14	7002.5	71.9	1528	US-09-939-853A-88	Sequence 88, Appli
15	6068.5	62.3	1303	US-10-618-281-63	Sequence 63, Appli
16	4697.5	48.3	1875	US-10-665-283-5	Sequence 5, Appli
17	4697.5	48.3	1911	US-10-665-283-2	Sequence 2, Appli
18	4667.5	48.0	1895	US-10-665-283-9	Sequence 9, Appli
19	4667.5	48.0	1901	US-10-665-283-10	Sequence 10, Appli
20	4491.5	46.1	1527	US-09-939-853A-83	Sequence 83, Appli
21	4491.5	46.1	1527	US-09-939-853A-83	Sequence 84, Appli
22	4491.5	46.1	1527	US-10-235-027-1342	Sequence 1342, Ap
23	4487.5	46.1	1527	US-10-888-503-6	Sequence 6, Appli
24	4487.5	46.1	1530	US-10-888-503-33	Sequence 33, Appli
25	4479.5	45.2	1530	US-10-618-281-55	Sequence 55, Appli
26	4396	45.2	1522	US-09-939-853A-85	Sequence 85, Appli
27	3966	40.7	1905	US-10-665-283-3	Sequence 3, Appli

28	3966	40.7	1941	5	US-10-665-283-12	Sequence 12, Appli
29	3936	40.4	1961	5	US-10-665-283-7	Sequence 7, Appli
30	3879.5	39.9	1548	5	US-10-807-466-28	Sequence 28, Appli
31	3860.5	39.7	1548	4	US-10-667-891-1	Sequence 1, Appli
32	3860.5	39.7	1548	5	US-10-807-466-1	Sequence 1, Appli
33	3846	39.5	1538	5	US-10-807-466-23	Sequence 23, Appli
34	3841	39.5	1538	5	US-10-807-466-23	Sequence 25, Appli
35	3835	39.4	1538	5	US-10-807-466-24	Sequence 24, Appli
36	3817.5	39.2	1539	5	US-10-807-466-26	Sequence 26, Appli
37	3817	39.2	1591	6	US-11-097-143-13266	Sequence 13266, A
38	3797	39.0	1538	5	US-10-807-466-27	Sequence 27, Appli
39	3779	38.8	1538	5	US-10-807-466-16	Sequence 16, Appli
40	3774	38.8	1538	5	US-10-807-466-18	Sequence 18, Appli
41	3773	38.8	1538	5	US-10-807-466-22	Sequence 22, Appli
42	3768	38.7	1538	5	US-10-807-466-17	Sequence 17, Appli
43	3754	38.6	1538	5	US-10-807-466-19	Sequence 19, Appli
44	3750.5	38.5	1539	5	US-10-807-466-20	Sequence 20, Appli
45	3733	38.4	1547	6	US-11-097-143-25248	Sequence 25248, A

ALIGNMENTS

RESULT 1									
US-10-665-283-4									
Sequence 4, Application US/10665283									
Publication No. US20050063989A1									
GENERAL INFORMATION:									
APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE									
APPLICANT: DERAND, Renaud									
APPLICANT: GARCIA, Elisabeth									
APPLICANT: PROST, Anne-Lise									
APPLICANT: REVILLAUD, Jean									
APPLICANT: VIVAUDOU, Michel									
TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF									
FILE REFERENCE: P263US104									
CURRENT APPLICATION NUMBER: US/10/665,283									
CURRENT FILING DATE: 2003-09-22									
NUMBER OF SEQ ID NOS: 25									
SOFTWARE: PatentIn version 3.1									
SEQ ID NO 4									
LENGTH: 1891									
TYPE: PRT									
ORGANISM: Homo sapiens									
US-10-665-283-4									
Query Match									
Best Local Similarity 100.0%; Score 9734; DB 5; Length 1891;									
Matches 1891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MALRGFCADGSDPLMDMNVNTWNTSNPDKCFQNTLVVWPCFYLMACFPFYLSRH	60						
DB	1	MALRGFCADGSDPLMDMNVNTWNTSNPDKCFQNTLVVWPCFYLMACFPFYLSRH	60						
QY	61	DRGIQMTPLNKTKTALGFLMTVCADLRYFWEBSRGFLAPVPLVSTLIGITLLA	120						
DB	61	DRGIQMTPLNKTKTALGFLMTVCADLRYFWEBSRGFLAPVPLVSTLIGITLLA	120						
QY	121	TFLLQERRRGVSSGIMLTFMLVALYCALAIRSKIMTLAKEDAQVDLFRDITFYVFS	180						
DB	121	TFLLQERRRGVSSGIMLTFMLVALYCALAIRSKIMTLAKEDAQVDLFRDITFYVFS	180						
QY	181	LLLIQVLSCFSDRSPLFSETIHDPNCPBSSASFISRIIFWMTGLIVRGYQPLEGSD	240						
DB	181	LLLIQVLSCFSDRSPLFSETIHDPNCPBSSASFISRIIFWMTGLIVRGYQPLEGSD	240						
QY	241	LWLNKEDTSEQVVPVLYKMKKECAKTRQPKVYVSSQDPAPQKSSSVYDANEVEAL	300						
DB	241	LWLNKEDTSEQVVPVLYKMKKECAKTRQPKVYVSSQDPAPQKSSSVYDANEVEAL	300						
QY	301	IVKSPQKENVNPLSKVLYKTFGPFYFLMSFFPKAIHDLAMFSGQILKLIFVNDTKAPD	360						
DB	301	IVKSPQKENVNPLSKVLYKTFGPFYFLMSFFPKAIHDLAMFSGQILKLIFVNDTKAPD	360						

QY 361 MGGYFTYVLLFVTAQLOTLVLHQYFHCFSVSGMRKTAIVGAVYRKALVITNSARKSTV 420
DB 361 MGGYFTYVLLFVTAQLOTLVLHQYFHCFSVSGMRKTAIVGAVYRKALVITNSARKSTV 420
QY 421 GEIYNLMSVDAQREMDLATYINMISAPLOYITLALYLMNLGSGVLAGVAVMLAMPVN 480
DB 421 GEIYNLMSVDAQREMDLATYINMISAPLOYITLALYLMNLGSGVLAGVAVMLAMPVN 480
QY 481 AVAMAKTKTYOVAMHKSNDRIKLMNEELINGIKVLKLYAMELAFDKYLAIROBELKYLK 540
DB 481 AVAMAKTKTYOVAMHKSNDRIKLMNEELINGIKVLKLYAMELAFDKYLAIROBELKYLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTIDENNIIDQAOTAFVSLAFNLIRPPLNLP 600
DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTIDENNIIDQAOTAFVSLAFNLIRPPLNLP 600
QY 601 MVISIYQASVSLKRLRFLSHELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660
DB 601 MVISIYQASVSLKRLRFLSHELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660
QY 661 LINGITPSPREGALVAVVQVCGKSSLLSALLAEMDKVGHVAIKGSAVYVPOQAMIOND 720
DB 661 LINGITPSPREGALVAVVQVCGKSSLLSALLAEMDKVGHVAIKGSAVYVPOQAMIOND 720
QY 721 SLRENILFEGCQLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKGVNLGGQKQVSLAR 780
DB 721 SLRENILFEGCQLEBPYRSVIOACALLPDLLEILPSGDRTEIGEKGVNLGGQKQVSLAR 780
QY 781 AVYENADYYLFDDPLSANDAVGKHIFENVIGPKMKLNKTRILVTHSMSTLPDYVITIV 840
DB 781 AVYENADYYLFDDPLSANDAVGKHIFENVIGPKMKLNKTRILVTHSMSTLPDYVITIV 840
QY 841 MSGGKISEMGSYOEILLARDGAFAEFLRTYASTEOQDAEENGCVGSGPKGKAKEMNGM 900
DB 841 MSGGKISEMGSYOEILLARDGAFAEFLRTYASTEOQDAEENGCVGSGPKGKAKEMNGM 900
QY 901 LVTDSAGKOLOROLSSSSSSSGDISSRHNSTAELOKAKKEETKLMKLEADKAOQTGYKL 960
DB 901 LVTDSAGKOLOROLSSSSSSSGDISSRHNSTAELOKAKKEETKLMKLEADKAOQTGYKL 960
QY 961 SVYWDYKAKIQLFISFLSIFLPMCHNVASALSNVLSMTDDPIYNGOETHKVLSTYGG 1020
DB 961 SVYWDYKAKIQLFISFLSIFLPMCHNVASALSNVLSMTDDPIYNGOETHKVLSTYGG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGLILASRCHVDLHSILSPMSFEERTSGMLVNRFSKEL 1080
DB 1021 ALGISOGIAVFGYSMAVSIIGLILASRCHVDLHSILSPMSFEERTSGMLVNRFSKEL 1080
QY 1081 DTVDSMIPEVIKMFMSLFFNYIGACTIVILLATPIAIIIPPLGLIYFFVQRFYVASSRQL 1140
DB 1081 DTVDSMIPEVIKMFMSLFFNYIGACTIVILLATPIAIIIPPLGLIYFFVQRFYVASSRQL 1140
QY 1141 KRLBSVSRSPYVSHFNETLLGVSVIRAFEBQERFHQSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLBSVSRSPYVSHFNETLLGVSVIRAFEBQERFHQSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLCVCVGCIALFPALEFVIRSHSLSAGLVGSLSVYSIQVTTYIYMLVMSSEMTNTIVA 1260
DB 1201 VRLCVCVGCIALFPALEFVIRSHSLSAGLVGSLSVYSIQVTTYIYMLVMSSEMTNTIVA 1260
QY 1261 VERLKEVSETEKEAPWQIOETAPPSMPQVGVVEFRNTCLARYEDLDVLRHINVTINGG 1320
DB 1261 VERLKEVSETEKEAPWQIOETAPPSMPQVGVVEFRNTCLARKREDDVDVLRHINVTINGG 1320
QY 1321 EKVGIVGRTGAGKSSLLGLFRINESAGEIIIDGINIAKIGLHDLRFKRTITIIPODPVLF 1380
DB 1321 EKVGIVGRTGAGKSSLLGLFRINESAGEIIIDGINIAKIGLHDLRFKRTITIIPODPVLF 1380
QY 1381 SGSLRMLNDPFSQYSDERVTWSTLELAHLKDPVSAIPDLKDHCCAGGENTLSTGQOQVYCL 1440
DB 1381 SGSLRMLNDPFSQYSDERVTWSTLELAHLKDPVSAIPDLKDHCCAGGENTLSTGQOQVYCL 1440

QY 1441 ARALLRRTKILVUDEAPNAVDELFTDDLIQSTIRFOBEDCVYLTIAHRLNTIMDYTRYIVL 1500
DB 1441 ARALLRRTKILVUDEAPNAVDELFTDDLIQSTIRFOBEDCVYLTIAHRLNTIMDYTRYIVL 1500
QY 1501 DKGEIOEYGAPSDILQORGLFYSMAKDAGLVGGGGGMLSRKGIIPREYVLTRLAEPDAE 1560
DB 1501 DKGEIOEYGAPSDILQORGLFYSMAKDAGLVGGGGGMLSRKGIIPREYVLTRLAEPDAE 1560
QY 1561 PRYTRRRRARFVSKKNCNVAHKNIREQGRFLQDVFTTVLDLKMPTHLLIFMSFLCSW 1620
DB 1561 PRYTRRRRARFVSKKNCNVAHKNIREQGRFLQDVFTTVLDLKMPTHLLIFMSFLCSW 1620
QY 1621 LLFAMVWMLIAFAHGDIAPEGINVCVTSIHSFSSAFLESIEVQVITIGFGGNVTEBCP 1680
DB 1621 LLFAMVWMLIAFAHGDIAPEGINVCVTSIHSFSSAFLESIEVQVITIGFGGNVTEBCP 1680
QY 1681 LALILIVQNTVGMINAMIGCIEMKTAQAHRAEFLIPSKHAVITLRHGLCFMRLVYG 1740
DB 1681 LALILIVQNTVGMINAMIGCIEMKTAQAHRAEFLIPSKHAVITLRHGLCFMRLVYG 1740
QY 1741 DLKRSMTIISATIHQVVRKTTSPGEVVPPLHQVDIPMENGVGNGIFLVAPLIIYHVIDS 1800
DB 1741 DLKRSMTIISATIHQVVRKTTSPGEVVPPLHQVDIPMENGVGNGIFLVAPLIIYHVIDS 1800
QY 1801 NSPLYDLAPSDLHHODLEIIVIEGVETTGITTOARTSYLADBIIMGQRFVPIVAEED 1860
DB 1801 NSPLYDLAPSDLHHODLEIIVIEGVETTGITTOARTSYLADBIIMGQRFVPIVAEED 1860
QY 1861 GRYSVDYSKFGNTIKVPLCTARQLEDERS 1891
DB 1861 GRYSVDYSKFGNTIKVPLCTARQLEDERS 1891

RESULT 2

US-10-665-283-1
; Sequence 1, Application US/10665283
; Publication No. US2005063989A1
; GENERAL INFORMATION:
; APPLICANT: COMMISSARIAT A L' ENERGIE ATOMIQUE
; APPLICANT: DERAND, Renaud
; APPLICANT: GARCIA, Elisabeth
; APPLICANT: PROST, Anne-Lise
; APPLICANT: REVILLOU, Jean
; APPLICANT: VIVAUDOU, Michel
; TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF
; FILE REFERENCE: P263US104
; CURRENT APPLICATION NUMBER: US/10/665, 283
; CURRENT FILING DATE: 2003-09-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1927
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-665-283-1

Query Match 100.0%; Score 9734; DB 5; Length 1927;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCGADSDPLMDMNVNTMNTSNDPFTKCFONTVLVWVPCFYLMACFPFFYLYLSRH 60
DB 1 MALRGFCGADSDPLMDMNVNTMNTSNDPFTKCFONTVLVWVPCFYLMACFPFFYLYLSRH 60
QY 61 DRGIYQMTPLNKTXTALGFLIMIVCMADLFYSFMSERSGRTIAPVFLVSPTLIGITTTLLA 120
DB 61 DRGIYQMTPLNKTXTALGFLIMIVCMADLFYSFMSERSGRTIAPVFLVSPTLIGITTTLLA 120
QY 121 TFLIQLEBRKGVQSSGIMLTFMLVALVICALIILSKITAKEDPAQVQVDFPDITFYVYFS 180
DB 121 TFLIQLEBRKGVQSSGIMLTFMLVALVICALIILSKITAKEDPAQVQVDFPDITFYVYFS 180
QY 181 LLLIQLVLSCFSDRSPLFSETIHDNPNCPBSSASFLSRITPFWMITGLIVRGYROPLEGSD 240

Db 181 LLLIOLVSCFSDSPLESETHIDENPCRESSASPLSRITFWMTGLIVRGYQGLESD 240
Qy 241 LMSLNKETSPQVAVVAVKMKKECAKTRKQPVKVVSSKDPAPQKSSKVDANEVVAL 300
Db 241 LMSLNKETSPQVAVVAVKMKKECAKTRKQPVKVVSSKDPAPQKSSKVDANEVVAL 300
Qy 301 IVKSPQKEMNPSLFVLYKTFGPYFLMSFFPKAHDLMFSGPOLIKLLIFVNDTKAPD 360
Db 301 IVKSPQKEMNPSLFVLYKTFGPYFLMSFFPKAHDLMFSGPOLIKLLIFVNDTKAPD 360
Qy 361 WQGFYVYLLFVTACLOTLVHQYFHCFSVGMRIKTAIVGAVYRKALVITNSARKSSTV 420
Db 361 WQGFYVYLLFVTACLOTLVHQYFHCFSVGMRIKTAIVGAVYRKALVITNSARKSSTV 420
Qy 421 GEIVNLSVDAQRFMDLATYINMTWSAPLOVILALYLLMLNGSVLGVAVMVLMPVN 480
Db 421 GEIVNLSVDAQRFMDLATYINMTWSAPLOVILALYLLMLNGSVLGVAVMVLMPVN 480
Qy 481 AVAMKTKTYOVAAHMKSKDNRIKLMNEILNGIKVLKLYAMELAFKDKYLAIRQBELKYLK 540
Db 481 AVAMKTKTYOVAAHMKSKDNRIKLMNEILNGIKVLKLYAMELAFKDKYLAIRQBELKYLK 540
Qy 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVITDENNIIDAQTAFYSLAFNILRFPPLNLP 600
Db 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVITDENNIIDAQTAFYSLAFNILRFPPLNLP 600
Qy 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPYKQGGGNTSTVRNATTTMARSDPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPYKQGGGNTSTVRNATTTMARSDPT 660
Qy 661 LINGTFSTPEGALVAVVGVGCGKSLLSALLAEMDKYGHVAIKGSVAVYPOQAMOND 720
Db 661 LINGTFSTPEGALVAVVGVGCGKSLLSALLAEMDKYGHVAIKGSVAVYPOQAMOND 720
Qy 721 SLRENIILFGCOLLEBPYRSVIOACALLPDEILPBGDRTEIGEGVNLSSGQOKQVSLAR 780
Db 721 SLRENIILFGCOLLEBPYRSVIOACALLPDEILPBGDRTEIGEGVNLSSGQOKQVSLAR 780
Qy 781 AVYSNADIYLFDDPLSAVDHAVGHIPENVIGPKMLKNKTRILVTHSMYSLPOVDVITV 840
Db 781 AVYSNADIYLFDDPLSAVDHAVGHIPENVIGPKMLKNKTRILVTHSMYSLPOVDVITV 840
Qy 841 MSGGKISEMGYOELLARDGAFAPLRTYASTBOGDAENGVYGVSGPGKEAQMENG 900
Db 841 MSGGKISEMGYOELLARDGAFAPLRTYASTBOGDAENGVYGVSGPGKEAQMENG 900
Qy 901 LVTDSAGKOLQROSLSSSSYSYSGDISRHNSYAELOKAEKKEETWKLMEADKAOTGOVKL 960
Db 901 LVTDSAGKOLQROSLSSSSYSYSGDISRHNSYAELOKAEKKEETWKLMEADKAOTGOVKL 960
Qy 961 SVYWDYKAIAGLFTSPLSIFLFCMNVHSAALSNYLSMTDPIVNGTOEHTKRLSVYG 1020
Db 961 SVYWDYKAIAGLFTSPLSIFLFCMNVHSAALSNYLSMTDPIVNGTOEHTKRLSVYG 1020
Qy 1021 ALGISOGIAGVGYMAVSIIGIILASRCLHVDLHLSILSPMSFFERTPSGULVNRFSKEL 1080
Db 1021 ALGISOGIAGVGYMAVSIIGIILASRCLHVDLHLSILSPMSFFERTPSGULVNRFSKEL 1080
Qy 1081 DTVDSMTPEVIKMGSLFENVIGACIVILLATPIAIIIPPLGLIYFVQRFYVASSROL 1140
Db 1081 DTVDSMTPEVIKMGSLFENVIGACIVILLATPIAIIIPPLGLIYFVQRFYVASSROL 1140
Qy 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
Db 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
Qy 1201 VRLCEVNCIYLFALPAVITSRHSLSAGLVSVSYSLQVTTYLNMLVRRMSSEMETNIVA 1260
Db 1201 VRLCEVNCIYLFALPAVITSRHSLSAGLVSVSYSLQVTTYLNMLVRRMSSEMETNIVA 1260
Qy 1261 VERLKEYSETEKAPMOIOETAPPSWPQVGRVEFRNYCLARYEDLDFVLRIHIVTNGG 1320
Db 1261 VERLKEYSETEKAPMOIOETAPPSWPQVGRVEFRNYCLARYEDLDFVLRIHIVTNGG 1320

Db 1261 VERLKEYSETEKAPMOIOETAPPSWPQVGRVEFRNYCLARYEDLDFVLRIHIVTNGG 1320
Qy 1321 EKVGIVERTGAGKSLTGLPRINESAGEIINDIGINAKIGHDLAEKTIITIODPLTF 1380
Db 1321 EKVGIVERTGAGKSLTGLPRINESAGEIINDIGINAKIGHDLAEKTIITIODPLTF 1380
Qy 1381 SGSLRMLDPPSQSDSEWTSLELAHKDFVSALPKLDHECAGEENLSVQOROLVCL 1440
Db 1381 SGSLRMLDPPSQSDSEWTSLELAHKDFVSALPKLDHECAGEENLSVQOROLVCL 1440
Qy 1441 ARALLRKTILVIDEATAVADLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRYVL 1500
Db 1441 ARALLRKTILVIDEATAVADLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRYVL 1500
Qy 1501 DKGEIOEYGAPSDILQORGLFYSAKADAGVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
Db 1501 DKGEIOEYGAPSDILQORGLFYSAKADAGVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
Qy 1561 PRYTRRRRARFVSKGNCNVAHKNIREQGRFLQDVFTLLVDLKMPTLILFTMSPLCSW 1620
Db 1561 PRYTRRRRARFVSKGNCNVAHKNIREQGRFLQDVFTLLVDLKMPTLILFTMSPLCSW 1620
Qy 1621 LLFAMVWMLIAFAHGLDAPGEGTVPCVTSIHSFSAFLRSIEVQVITIGFGRMVTBEC 1680
Db 1621 LLFAMVWMLIAFAHGLDAPGEGTVPCVTSIHSFSAFLRSIEVQVITIGFGRMVTBEC 1680
Qy 1681 LAIILILVONIVGMINAIVMGLCTFMKTAQAHRAEFLISKAAVITLRHGRCLCPMLRVG 1740
Db 1681 LAIILILVONIVGMINAIVMGLCTFMKTAQAHRAEFLISKAAVITLRHGRCLCPMLRVG 1740
Qy 1741 DLKSMIISATIHQVVRKTTSPGEVVPVPHQVDIPENGVGGGIFLVAPLIYHVIDS 1800
Db 1741 DLKSMIISATIHQVVRKTTSPGEVVPVPHQVDIPENGVGGGIFLVAPLIYHVIDS 1800
Qy 1801 NSPLVLDAPSDLHMHODLEIIVILEGVETTGITTOARTSYLADEIIMGORFVPIVAED 1860
Db 1801 NSPLVLDAPSDLHMHODLEIIVILEGVETTGITTOARTSYLADEIIMGORFVPIVAED 1860
Qy 1861 GRYSVDYSKFGNTIKVPTPLCTARQLEDDRS 1891
Db 1861 GRYSVDYSKFGNTIKVPTPLCTARQLEDDRS 1891

RESULT 3
US-10-665-283-6
; Sequence 6, Application US/10665283
; Publication No. US20050063989A1
; GENERAL INFORMATION:
; APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
; APPLICANT: DERAND, Renaud
; APPLICANT: GARCIA, Elisabeth
; APPLICANT: PROST, Anne-Lise
; APPLICANT: REVILLAUD, Jean
; APPLICANT: VIVAUDOU, Michel
; TITLE OR INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREO
; FILE REFERENCE: P26305104
; CURRENT APPLICATION NUMBER: US/10/665,283
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1927
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-665-283-6

Query Match 100.0%; Score 9734; DB 5; Length 1927;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRGFCSDGSDPLMDMNTVNTSNDPTKCFONTVLVWVPCFYLMACPPFYFLYSRH 60
Db 1 MALRGFCSDGSDPLMDMNTVNTSNDPTKCFONTVLVWVPCFYLMACPPFYFLYSRH 60

QY 61 DRGYIOWTPLNKKTALGFLIMIVCMADLFYSFEMERSKGIPLAPVFLVSPILLGITTLA 120
Db 61 DRGYIOWTPLNKKTALGFLIMIVCMADLFYSFEMERSKGIPLAPVFLVSPILLGITTLA 120
QY 121 TFLIOLEBRRKVOSSGIMLTFWLVALVCAALILSKINTALKEADAVLDFDITFYVFS 180
Db 121 TFLIOLEBRRKVOSSGIMLTFWLVALVCAALILSKINTALKEADAVLDFDITFYVFS 180
QY 181 LLLIQLVASCSDSPLESETIHDNPPCBESSASLRIITFMWITGLIVRGROPLESSD 240
Db 181 LLLIQLVASCSDSPLESETIHDNPPCBESSASLRIITFMWITGLIVRGROPLESSD 240
QY 241 LMSLINKEDTSEOVVPLVKNMKKECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVVAL 300
Db 241 LMSLINKEDTSEOVVPLVKNMKKECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVVAL 300
QY 301 IVKSPCKEMNPSPLEKVTGTPFTLMSFFPKAIDHLMFSGPOLIKLIFVNDTKAPD 360
Db 301 IVKSPCKEMNPSPLEKVTGTPFTLMSFFPKAIDHLMFSGPOLIKLIFVNDTKAPD 360
QY 361 WQGFYTYTLFVTAQLOTLVHOVPHICFVSGMRKTPAVIGAVYRKALVITNSARKSTV 420
Db 361 WQGFYTYTLFVTAQLOTLVHOVPHICFVSGMRKTPAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVILALYLLMLNGPSVLAVGAVMVLMPVN 480
Db 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVILALYLLMLNGPSVLAVGAVMVLMPVN 480
QY 481 AVMAKTKTYOVAMHKSQONRIKLMNELNGIKVLKLYAMELAFKDXLAIROBELKYLK 540
Db 481 AVMAKTKTYOVAMHKSQONRIKLMNELNGIKVLKLYAMELAFKDXLAIROBELKYLK 540
QY 541 KSAVLASNGTWTWCTPLVALCTPAVYVTDENNILDAQTAFSALFNILRPDLNLP 600
Db 541 KSAVLASNGTWTWCTPLVALCTPAVYVTDENNILDAQTAFSALFNILRPDLNLP 600
QY 601 MVISSIVQASVSLKRLIFLSHEBELPDSIERRPYKDGCTNSITVRNATFTWASDPPT 660
Db 601 MVISSIVQASVSLKRLIFLSHEBELPDSIERRPYKDGCTNSITVRNATFTWASDPPT 660
QY 661 LNGTHFSPREGALVAVGVQVCGKSSLLSALLAENDKVBGHVAKGSVAVYVPOQAMIOND 720
Db 661 LNGTHFSPREGALVAVGVQVCGKSSLLSALLAENDKVBGHVAKGSVAVYVPOQAMIOND 720
QY 721 SLRENIIFGCOLLEBYRSVIOACALLPDLLETLPBGDTEIGEKVNSGQKQVSLAR 780
Db 721 SLRENIIFGCOLLEBYRSVIOACALLPDLLETLPBGDTEIGEKVNSGQKQVSLAR 780
QY 781 AVYSNADIYLPDDPLSAVDAHVKHIFENVIGPKGMLKNKTRILVTHSMSYLPQVDVILV 840
Db 781 AVYSNADIYLPDDPLSAVDAHVKHIFENVIGPKGMLKNKTRILVTHSMSYLPQVDVILV 840
QY 841 MSGGKISMSGSYOELLARDGAFAEFLRTYASTEOBDAEENGTVSGSPGKEAKEMNGM 900
Db 841 MSGGKISMSGSYOELLARDGAFAEFLRTYASTEOBDAEENGTVSGSPGKEAKEMNGM 900
QY 901 LVTBSAGQOLOROLSSSSSSSGDISRHNSSTAELQKAKKKEBTMKLMEADAOQOYKL 960
Db 901 LVTBSAGQOLOROLSSSSSSSGDISRHNSSTAELQKAKKKEBTMKLMEADAOQOYKL 960
QY 961 SVYMDYMAKIGLFIIFLSIFLPMCHVASALASNYMLMTDPIYNGOEHKVALSYG 1020
Db 961 SVYMDYMAKIGLFIIFLSIFLPMCHVASALASNYMLMTDPIYNGOEHKVALSYG 1020
QY 1021 ALGISOGIAVFGYMAVSIIGILLASRCLHVDLHSLRSPMSFFERTSPGNLVNFSKEL 1080
Db 1021 ALGISOGIAVFGYMAVSIIGILLASRCLHVDLHSLRSPMSFFERTSPGNLVNFSKEL 1080
QY 1081 DTUUSMIEBVIMFNGSLPNVIGACIVILATPIAIIIPBLGILYFPQRPYVASSROL 1140
Db 1081 DTUUSMIEBVIMFNGSLPNVIGACIVILATPIAIIIPBLGILYFPQRPYVASSROL 1140

QY 1141 KRLSVSRSPVYSHFNFTLLGVSVITRAPEOEERFIHOSDLKVDENQAYYPSIYANRWLA 1200
Db 1141 KRLSVSRSPVYSHFNFTLLGVSVITRAPEOEERFIHOSDLKVDENQAYYPSIYANRWLA 1200
QY 1201 VRLCEVNCIVLPALPAVLSRHSLSAGLVASYSLOVTTYANMLVRMSSEMETNIVA 1260
Db 1201 VRLCEVNCIVLPALPAVLSRHSLSAGLVASYSLOVTTYANMLVRMSSEMETNIVA 1260
QY 1261 VERLEKSETEKEAPMOIOETAPSSWPOVGRVFRNYCLRYREDDLVFLSHINVTINGG 1320
Db 1261 VERLEKSETEKEAPMOIOETAPSSWPOVGRVFRNYCLRYREDDLVFLSHINVTINGG 1320
QY 1321 EKVGIQVBTGAKSSLTFLGLFRINESABGEIIDGINIAKIGLHDLRPKTTIIQDPVLF 1380
Db 1321 EKVGIQVBTGAKSSLTFLGLFRINESABGEIIDGINIAKIGLHDLRPKTTIIQDPVLF 1380
QY 1381 SGSLRMLNDPSSQVSDSEWMTSLFLAKOPVSLPKLDBECAGGGENLSVQOROLVCL 1440
Db 1381 SGSLRMLNDPSSQVSDSEWMTSLFLAKOPVSLPKLDBECAGGGENLSVQOROLVCL 1440
QY 1441 ARALLRKTILVDEATRAVDELTDLLIOGTIRTOFEDCTVLTIAHRLNTIMDYTRVVL 1500
Db 1441 ARALLRKTILVDEATRAVDELTDLLIOGTIRTOFEDCTVLTIAHRLNTIMDYTRVVL 1500
QY 1501 DKGEIQEYGAAPSDLLQORGLFYSSAKDAGLVGGGGGMLSRKGIIPBEYVLTILAEDPAE 1560
Db 1501 DKGEIQEYGAAPSDLLQORGLFYSSAKDAGLVGGGGGMLSRKGIIPBEYVLTILAEDPAE 1560
QY 1561 PRYTRERRARFVSKKNCVNAHKNIREOGFLODVEVTLVDTLKPHTLLIFMSFLCSW 1620
Db 1561 PRYTRERRARFVSKKNCVNAHKNIREOGFLODVEVTLVDTLKPHTLLIFMSFLCSW 1620
QY 1621 LFPAMWMLIFAFAGDLPAGEGTNVCVTSIHSFSSAFLFSIEVQVITIGFGRWTEBCEP 1680
Db 1621 LFPAMWMLIFAFAGDLPAGEGTNVCVTSIHSFSSAFLFSIEVQVITIGFGRWTEBCEP 1680
QY 1681 LAILILIVONIVGIMINAMIGCIEMKTAQHRRPAETLIFSKHAVITLRHRLCPMLRVG 1740
Db 1681 LAILILIVONIVGIMINAMIGCIEMKTAQHRRPAETLIFSKHAVITLRHRLCPMLRVG 1740
QY 1741 DLKRSMTIISATIHQVARKTSPGEVAVPLHQVDIPMENGVGNGIFLVAPLLIVHVIDS 1800
Db 1741 DLKRSMTIISATIHQVARKTSPGEVAVPLHQVDIPMENGVGNGIFLVAPLLIVHVIDS 1800
QY 1801 NSPLVDLAPSDLHHODLEIIVILEGVETGITTQARTSYLADEILMGOBFVIAEED 1860
Db 1801 NSPLVDLAPSDLHHODLEIIVILEGVETGITTQARTSYLADEILMGOBFVIAEED 1860
QY 1861 GRYSVDYSKFGNTIKVPTPLCTARQULDEDRS 1891
Db 1861 GRYSVDYSKFGNTIKVPTPLCTARQULDEDRS 1891

RESULT 4
US-10-665-283-8
; Sequence 8, Application US/1065283
; Publication No. US2005063989A1
; GENERAL INFORMATION:
; APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
; APPLICANT: DERAND, Renaud
; APPLICANT: GARCIA, Elisabeth
; APPLICANT: PROST, Anne-Lise
; APPLICANT: REVILLOUP, Jean
; APPLICANT: VIANOUOU, Michel
; TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF
; FILE REFERENCE: F263US104
; CURRENT APPLICATION NUMBER: US/10/665, 283
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 1947
; TYPE: PRT

ORGANISM: Homo sapiens
US-10-665-283-8

Query Match 99.7%; Score 9704; DB 5; Length 1947;
Beet Local Similarity 99.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 20; Gaps 2;

1 MALRFGCADSDPMDMNVMTNTSNPDFTKCFONTLVVWVPCFLMACFPPEFLYLSRH 60
1 MALRFGCADSDPMDMNVMTNTSNPDFTKCFONTLVVWVPCFLMACFPPEFLYLSRH 60
61 DRGYIOMTPLNKTALGFLMIVCMADLFYSFWERSRGIFLAPFVLSPTLLGTTLLA 120
61 DRGYIOMTPLNKTALGFLMIVCMADLFYSFWERSRGIFLAPFVLSPTLLGTTLLA 120
121 TFLIOLERRKGVOSGIMLTWLVALVCAALIRSKIMTALKEDAOVDLPFDITPYVFS 180
121 TFLIOLERRKGVOSGIMLTWLVALVCAALIRSKIMTALKEDAOVDLPFDITPYVFS 180
181 LLLIQLVASCDSRDLSETHDNPCESSASLSRITFMWITGLIVRGYRQPLEGSD 240
181 LLLIQLVASCDSRDLSETHDNPCESSASLSRITFMWITGLIVRGYRQPLEGSD 240
241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKOPVKVYVSSKDPAPCKESSKYDANEVEAL 300
241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKOPVKVYVSSKDPAPCKESSKYDANEVEAL 300
301 IVKSPQKEMNPDLFEVLYKTTGPPYFLMSFFPKAHDLMFSGPOLIKLIRVNDTKAPD 360
301 IVKSPQKEMNPDLFEVLYKTTGPPYFLMSFFPKAHDLMFSGPOLIKLIRVNDTKAPD 360
361 WQGFYVTLFVTCLOTLVHOVPHLCFVSGMRKKTAVIGAVYRKALVITNSARKSTV 420
361 WQGFYVTLFVTCLOTLVHOVPHLCFVSGMRKKTAVIGAVYRKALVITNSARKSTV 420
421 GEIVNLSVDAQRFMDLATYINMISAPLOVILALYILMLNGSVLGVAVMLVAVN 480
421 GEIVNLSVDAQRFMDLATYINMISAPLOVILALYILMLNGSVLGVAVMLVAVN 480
481 AVMAKTKTYOVAVHMKSKDNRIKLMNEILNGIKVILKAYAMELAFKDYALIRQELKYLK 540
481 AVMAKTKTYOVAVHMKSKDNRIKLMNEILNGIKVILKAYAMELAFKDYALIRQELKYLK 540
541 KSAVLASNGFTWCTPFLVALCTPAVYVTTIDENNIIIDAQTAFLSALFNLIRPDLTP 600
541 KSAVLASNGFTWCTPFLVALCTPAVYVTTIDENNIIIDAQTAFLSALFNLIRPDLTP 600
601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTMARSDPT 660
601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKDGGTNSITVRNATFTMARSDPT 660
661 LMGITFSPREGALVAVVGVCGKSSLLSALLAEMDKVGHVAKGSVAVVYPOQAMIOND 720
661 LMGITFSPREGALVAVVGVCGKSSLLSALLAEMDKVGHVAKGSVAVVYPOQAMIOND 720
721 SLRNNILFGCOLLEPPYRSVYQACALLPDLLEILPSGDTETIGEGVNLSSGQKQVSLAR 780
721 SLRNNILFGCOLLEPPYRSVYQACALLPDLLEILPSGDTETIGEGVNLSSGQKQVSLAR 780
781 AVYENADIYLPDDPLSAVDHAKHIFENVIGPKGMLKNKTRILVTHSMSTLPQVDYIIV 840
781 AVYENADIYLPDDPLSAVDHAKHIFENVIGPKGMLKNKTRILVTHSMSTLPQVDYIIV 840
841 MSGKISBMSGYOELLARDGAFAPFLRTYASTEOBDAEENGVTGVSQPKAKOMENGM 900
841 MSGKISBMSGYOELLARDGAFAPFLRTYASTEOBDAEENGVTGVSQPKAKOMENGM 900
901 LVTSAGKOLROISSSSSSSGDISRHNSSTAELOKAKAKEFTWTKMEADKAOTGOYKL 960
901 LVTSAGKOLROISSSSSSSGDISRHNSSTAELOKAKAKEFTWTKMEADKAOTGOYKL 960
961 SVYDYMKAIGLFTSLIFLFCMCHVASALASNYWLSMTDDPIVNGTOEHTKYRLSYVG 1020
961 SVYDYMKAIGLFTSLIFLFCMCHVASALASNYWLSMTDDPIVNGTOEHTKYRLSYVG 1020

961 SVYDYMKAIGLFTSLIFLFCMCHVASALASNYWLSMTDDPIVNGTOEHTKYRLSYVG 1020
1021 ALGISOGIANTVGYMAVNSIGGILASRCLHVDLHLSILRSPMSFFERTPSGULVNRFSKL 1080
1021 ALGISOGIANTVGYMAVNSIGGILASRCLHVDLHLSILRSPMSFFERTPSGULVNRFSKL 1080
1081 DTVDSMPEVIMKMGSLFNIVIGACIVILATPTAAIIIPPLGIIYFVQRFYVASSROL 1140
1081 DTVDSMPEVIMKMGSLFNIVIGACIVILATPTAAIIIPPLGIIYFVQRFYVASSROL 1140
1141 KRLESVSRSPYSHNETLGVSVIRAFEEQEREIHOSDLKYDENOKAYPSIYANRWLA 1200
1141 KRLESVSRSPYSHNETLGVSVIRAFEEQEREIHOSDLKYDENOKAYPSIYANRWLA 1200
1201 VRLCEVGNCTVLPALPAVIRSRHLSAGLVLSYSISQVTTYANLVYRMSSEMETNIVA 1260
1201 VRLCEVGNCTVLPALPAVIRSRHLSAGLVLSYSISQVTTYANLVYRMSSEMETNIVA 1260
1261 VERLKEYSETEKEAPMOIQETAPPSWPOVGRVFRNYCLARYEDLDPLVLRHINVTING 1320
1261 VERLKEYSETEKEAPMOIQETAPPSWPOVGRVFRNYCLARYEDLDPLVLRHINVTING 1320
1321 EKVGIVRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
1321 EKVGIVRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
1381 SGSLRPMULDPPSQYSDBEWTSLELAHKDFVSAALPKLDHECAEGGENLSVGOROLVCL 1440
1381 SGSLRPMULDPPSQYSDBEWTSLELAHKDFVSAALPKLDHECAEGGENLSVGOROLVCL 1440
1441 ABALLRKTILVDEATAVADLETDDLIQSTIRQFEDCTVLTJAHNRANTMDTRVVL 1500
1441 ABALLRKTILVDEATAVADLETDDLIQSTIRQFEDCTVLTJAHNRANTMDTRVVL 1500
1501 DKGEIOBYGAPSDLLQORGLFYSMAXDAGLVGGGGGMLSRKGIIPBEYVLTSLAEDPAE 1560
1501 DKGEIOBYGAPSDLLQORGLFYSMAXDAGLVGGGGGMLSRKGIIPBEYVLTSLAEDPAE 1560
1561 PRYTRERRARFVSKKNCNVAHKNIRBOGRFLODVFVTLVDLKWPHTLILFTMSFLCSW 1620
1561 PRYTRERRARFVSKKNCNVAHKNIRBOGRFLODVFVTLVDLKWPHTLILFTMSFLCSW 1620
1621 LLFAMVWMLTAFAG-----DLAP-----GEGTNVPCVTSIHSSSALFL 1660
1621 LLFAMVWMLTAFAGDLIYAMEKGIIDLAPPYDVPYAGGINVPCVTSIHSSSALFL 1660
1661 SIEVOVITGFGRAWTECEPLAILILIVONIVGIMINAMIGCIFMKTAAQHRRAETLIF 1720
1661 SIEVOVITGFGRAWTECEPLAILILIVONIVGIMINAMIGCIFMKTAAQHRRAETLIF 1720
1721 SKHAVITLRHGRCLCEMLRVGDLRKSMTISATIHQVVRKTTSPBGEVVPPLHQVDIPMENG 1780
1721 SKHAVITLRHGRCLCEMLRVGDLRKSMTISATIHQVVRKTTSPBGEVVPPLHQVDIPMENG 1780
1741 SKHAVITLRHGRCLCEMLRVGDLRKSMTISATIHQVVRKTTSPBGEVVPPLHQVDIPMENG 1800
1741 SKHAVITLRHGRCLCEMLRVGDLRKSMTISATIHQVVRKTTSPBGEVVPPLHQVDIPMENG 1800
1781 VGGNGIFLVAPLIIYHVIDNSPLXYDLAPBDLHHOOLEIIVILEGVEVETGTTQARTS 1840
1781 VGGNGIFLVAPLIIYHVIDNSPLXYDLAPBDLHHOOLEIIVILEGVEVETGTTQARTS 1840
1841 YLADIELMGORFVPIVAEEDGRYSVDYSKGGNTIKVPTPLCTARQLEDEDS 1891
1841 YLADIELMGORFVPIVAEEDGRYSVDYSKGGNTIKVPTPLCTARQLEDEDS 1891
1861 YLADIELMGORFVPIVAEEDGRYSVDYSKGGNTIKVPTPLCTARQLEDEDS 1911
1861 YLADIELMGORFVPIVAEEDGRYSVDYSKGGNTIKVPTPLCTARQLEDEDS 1911

RESULT 5
US-09-939-853A-86
; Sequence 86, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939, 853A
; CURRENT FILING DATE: 2001-08-27

PRIOR APPLICATION NUMBER: 60/228,191
PRIOR FILING DATE: 2000-08-25
PRIOR APPLICATION NUMBER: 60/267,300
PRIOR FILING DATE: 2001-02-08
PRIOR APPLICATION NUMBER: 60/269,961
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/277,337
PRIOR FILING DATE: 2001-03-20
NUMBER OF SEQ ID NOS: 159
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 86
LENGTH: 1531
TYPE: PRT
ORGANISM: Homo sapiens
US-09-939-853A-86

Query Match 80.7%; Score 7860; DB 3; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALGFCGADGSDPLMDNNTWNTSNPDFTKCFONTVLVWVCFYLMACFPFFLYLSRH 60
DB 1 MALGFCGADGSDPLMDNNTWNTSNPDFTKCFONTVLVWVCFYLMACFPFFLYLSRH 60
QY 61 DRGYIQTPLNKTXTALGFLIMVCMADLFYSFWERSRGIFLAPVFLVSPULLGITLLA 120
DB 61 DRGYIQTPLNKTXTALGFLIMVCMADLFYSFWERSRGIFLAPVFLVSPULLGITLLA 120
QY 121 TFLQLERKRGVQSSGIMLTFWLVLCALALRSKIMTALKEQVQLFRDITFYVFS 180
DB 121 TFLQLERKRGVQSSGIMLTFWLVLCALALRSKIMTALKEQVQLFRDITFYVFS 180
QY 181 LLLIQLVLSGSDRSPLFSETHDNPCESSASFLSRITTFWIMIGLIVRGYRPLEGSD 240
DB 181 LLLIQLVLSGSDRSPLFSETHDNPCESSASFLSRITTFWIMIGLIVRGYRPLEGSD 240
QY 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKQPVKVYSSKDPADQPKSSKVDANEVBAL 300
DB 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKQPVKVYSSKDPADQPKSSKVDANEVBAL 300
QY 301 IVKSPQKWNPSLFFVLYKTGPVFLMSFPFKAHDMFGSPQLIKLIKRVNDTKAPD 360
DB 301 IVKSPQKWNPSLFFVLYKTGPVFLMSFPFKAHDMFGSPQLIKLIKRVNDTKAPD 360
QY 361 WQGFYVTLVETACLOTLVHQYHICFVSGMRITKAVIGAVYKALVITNSAKSSTV 420
DB 361 WQGFYVTLVETACLOTLVHQYHICFVSGMRITKAVIGAVYKALVITNSAKSSTV 420
QY 421 GEIVNLSVDAQRFDLATYINMISAPLOVILALYLLMLNIGPSVLGAVAVMLMVEVN 480
DB 421 GEIVNLSVDAQRFDLATYINMISAPLOVILALYLLMLNIGPSVLGAVAVMLMVEVN 480
QY 481 AVMMKTKTYOVANHKSNDNRKLMNEILNGIKVLYAMELAFODKYLAIROEELKYLK 540
DB 481 AVMMKTKTYOVANHKSNDNRKLMNEILNGIKVLYAMELAFODKYLAIROEELKYLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTFPAVYTTIDENNILDAQZAFVSLALFNILRPNTILP 600
DB 541 KSAVLSAVGTFTWCTPFLVALCTFPAVYTTIDENNILDAQZAFVSLALFNILRPNTILP 600
QY 601 MVISIVQASVSLKRLIFLSHELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
DB 601 MVISIVQASVSLKRLIFLSHELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
QY 661 LNGTHFSIPEGALVAVGVQGVGKSSLSALLAENDKXEGHVAIGSAVAVVQQAAMIOND 720
DB 661 LNGTHFSIPEGALVAVGVQGVGKSSLSALLAENDKXEGHVAIGSAVAVVQQAAMIOND 720
QY 721 SLRENIILGCOLLEBYRYSVIOACALPLDLETLSGDRTEIEGKGNISGGOKGVSLAR 780
DB 721 SLRENIILGCOLLEBYRYSVIOACALPLDLETLSGDRTEIEGKGNISGGOKGVSLAR 780
QY 781 AVYSNADIYLFDDPLSAVDAAVGHKAI FENVIGPKGMKKNKTRILVTHSMSTLPQVDVILV 840

DB 781 AVYSNADIYLFDDPLSAVDAAVGHKAI FENVIGPKGMKKNKTRILVTHSMSTLPQVDVILV 840
QY 841 MSGGKISEMSGYOEILLARDGAFAEFLRTYASTBOEODAEENGVTGVSQPGKEALOMENGM 900
DB 841 MSGGKISEMSGYOEILLARDGAFAEFLRTYASTBOEODAEENGVTGVSQPGKEALOMENGM 900
QY 901 LVYDSAGKOLQOROLSSSSSYSGDISRHNSYAELOKAEKKEETWKLMEADKAOTGVYKL 960
DB 901 LVYDSAGKOLQOROLSSSSSYSGDISRHNSYAELOKAEKKEETWKLMEADKAOTGVYKL 960
QY 961 SVYWDYMKATGLFISFLSIFLPMCNHVSALASNYWLSLWDDPLVNGTOEHTKRLSYYG 1020
DB 961 SVYWDYMKATGLFISFLSIFLPMCNHVSALASNYWLSLWDDPLVNGTOEHTKRLSYYG 1020
QY 1021 ALGISOGIAVFGYMAVSIIGIILASRCILHVDLHSHIIRSPMSFEPRTSGULVNRFSKEL 1080
DB 1021 ALGISOGIAVFGYMAVSIIGIILASRCILHVDLHSHIIRSPMSFEPRTSGULVNRFSKEL 1080
QY 1081 DTVDSMIPEVIKMEGSLFNVIGACIYLATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
DB 1081 DTVDSMIPEVIKMEGSLFNVIGACIYLATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
QY 1141 KRLSVSRSPYSHRNFTLGVSYIRAEQERETIHOSDLKVDENOKAYYSIVANRWLA 1200
DB 1141 KRLSVSRSPYSHRNFTLGVSYIRAEQERETIHOSDLKVDENOKAYYSIVANRWLA 1200
QY 1201 VRLSCVNGCIYLPALPFAVISRHSLSAGLVGLSVSYSLQVTTYLNMVLRMSSEMETNIVA 1260
DB 1201 VRLSCVNGCIYLPALPFAVISRHSLSAGLVGLSVSYSLQVTTYLNMVLRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPWQOETAPSSWPQVGRVEFRNYCLRYRDLDFVLHINVTINGG 1320
DB 1261 VERLKEYSETEKAPWQOETAPSSWPQVGRVEFRNYCLRYRDLDFVLHINVTINGG 1320
QY 1321 EKVGVGTGAGKSLTGLFRINESAGEIILIDGINAKIGLHDLRKITIIIPDDPYLF 1380
DB 1321 EKVGVGTGAGKSLTGLFRINESAGEIILIDGINAKIGLHDLRKITIIIPDDPYLF 1380
QY 1381 SGLSRMLNDPFSQYSDDEEWTSLSLAHLKDVSAALPDLDBECAGENLSVGOROLVCL 1440
DB 1381 SGLSRMLNDPFSQYSDDEEWTSLSLAHLKDVSAALPDLDBECAGENLSVGOROLVCL 1440
QY 1441 ARALLRKTILVLEBATAVDLETTDILQSTIRTOFEDCTVLTIAHRLNTIMDYTRVILV 1500
DB 1441 ARALLRKTILVLEBATAVDLETTDILQSTIRTOFEDCTVLTIAHRLNTIMDYTRVILV 1500
QY 1501 DKGEIOEYGAPSDILQORGLFYMAKDAGLV 1531
DB 1501 DKGEIOEYGAPSDILQORGLFYMAKDAGLV 1531

RESULT 6
US-10-667-891-6
Sequence 6, Application US/10667891
Publication No. US20040171024A1
GENERAL INFORMATION:
APPLICANT: ROTH, CHARLES W.
APPLICANT: BREY, PAUL T.
APPLICANT: HOLM, INGE
APPLICANT: GRAILLES, MARINE
APPLICANT: RZHETSKY, ANDREY
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND TITLE OF INVENTION: ANOPHELES
FILE REFERENCE: 03495-0294-00000
CURRENT APPLICATION NUMBER: US/10/667,891
CURRENT FILING DATE: 2003-09-23
PRIOR APPLICATION NUMBER: 60/413,469
PRIOR FILING DATE: 2002-09-26
NUMBER OF SEQ ID NOS: 76
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO 6
LENGTH: 1531

TYPE: PRT
ORGANISM: Homo sapiens
US-10-667-891-6

Query Match 80.7%; Score 7860; DB 4; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSDGSDPLMDNMVNTWNTSNPDTKCFONTVLVWPCFYLMACPPFYFLYSRH 60
DB 1 MALRGFCSDGSDPLMDNMVNTWNTSNPDTKCFONTVLVWPCFYLMACPPFYFLYSRH 60

QY 61 DRGVIQMTPLNKTKTALGFLIMIVCMADLFYSFWMERSGFIPLAVFLVSPFLIGITLLA 120
DB 61 DRGVIQMTPLNKTKTALGFLIMIVCMADLFYSFWMERSGFIPLAVFLVSPFLIGITLLA 120

QY 121 TFLIOLERRKGVGSSGIMLTFMLVALVCAALILRSKIMTALKEDAOYDLFRDITFYVYFS 180
DB 121 TFLIOLERRKGVGSSGIMLTFMLVALVCAALILRSKIMTALKEDAOYDLFRDITFYVYFS 180

QY 181 LLLIQLVLSGSDSSPLFSETHIDNPNCPRESSASFISRTITWMTGLIVRGYRPLGSD 240
DB 181 LLLIQLVLSGSDSSPLFSETHIDNPNCPRESSASFISRTITWMTGLIVRGYRPLGSD 240

QY 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKOPVKNVSSKDPAPPKSSSKVDANEVYAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKOPVKNVSSKDPAPPKSSSKVDANEVYAL 300

QY 301 IVKSPQKEMNPSPFLKVLTKTGPYFLMSFPFKAIHDLMPFSGPOLIKLILKFNVDTRAPD 360
DB 301 IVKSPQKEMNPSPFLKVLTKTGPYFLMSFPFKAIHDLMPFSGPOLIKLILKFNVDTRAPD 360

QY 361 WQGFYFVLLFVTAQLOTLVHQYFHI CFVSGMRKTAIVGAYRKALVITNSARKSTV 420
DB 361 WQGFYFVLLFVTAQLOTLVHQYFHI CFVSGMRKTAIVGAYRKALVITNSARKSTV 420

QY 421 GEIYNLMSVDAORFMDLATYINMWSA PLQVYTLALYLLMINTGSLVAGVAVMVLAMPVN 480
DB 421 GEIYNLMSVDAORFMDLATYINMWSA PLQVYTLALYLLMINTGSLVAGVAVMVLAMPVN 480

QY 481 AVNAMKTKTYOVAMHMSKDNRIKLMNEILNGIKVLYLAMELAFKDKVLAIROBELKVLK 540
DB 481 AVNAMKTKTYOVAMHMSKDNRIKLMNEILNGIKVLYLAMELAFKDKVLAIROBELKVLK 540

QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITIDENNIIDAQTAVSLAFNILRPPLILP 600
DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITIDENNIIDAQTAVSLAFNILRPPLILP 600

QY 601 MVISSIVQASVSLKRLRIFLSHEELBPDSEIRRPVKDGGGNSITVNNATFTWARSDDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEELBPDSEIRRPVKDGGGNSITVNNATFTWARSDDPT 660

QY 661 LMGITFESIPEGALVAVVGVCCKSSLLSALLAEMDVKEGHVAIKGSVAVYPOQAWIOND 720
DB 661 LMGITFESIPEGALVAVVGVCCKSSLLSALLAEMDVKEGHVAIKGSVAVYPOQAWIOND 720

QY 721 SLRENIIFGCOLBEPYRYSVIOACALLPDLIELPSGRTEIGEXVNLSCGQKQVSLAR 780
DB 721 SLRENIIFGCOLBEPYRYSVIOACALLPDLIELPSGRTEIGEXVNLSCGQKQVSLAR 780

QY 781 AVYSNNDIYLPDPLSVDVAHVGHITFENYIGPKMKNTRILLVTHSMSYLPQVDVIIV 840
DB 781 AVYSNNDIYLPDPLSVDVAHVGHITFENYIGPKMKNTRILLVTHSMSYLPQVDVIIV 840

QY 841 MSGGKISEMGSYOELLARDAFAFLRTVASTBOEDABENGVTVSGPKGEAKOMENGM 900
DB 841 MSGGKISEMGSYOELLARDAFAFLRTVASTBOEDABENGVTVSGPKGEAKOMENGM 900

QY 901 LVNDSAGKQORQULSSSSSYSGDISRRHNSTAEIQAFAKKEEYWKLMENDKQOTGYVL 960
DB 901 LVNDSAGKQORQULSSSSSYSGDISRRHNSTAEIQAFAKKEEYWKLMENDKQOTGYVL 960

QY 961 SVYWDYKAIQGLFISLIFLMCNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020

DB 961 SVYWDYKAIQGLFISLIFLMCNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020

QY 1021 ALGISQIAVFGYMAVSIGIILASRCLHYDLHSILRSFMSFEPFPGNLVNRSEKEL 1080
DB 1021 ALGISQIAVFGYMAVSIGIILASRCLHYDLHSILRSFMSFEPFPGNLVNRSEKEL 1080

QY 1081 DTWDSMILEVINKPMGSLFNVGACIVTILATPAALIIIPPLIGLYFVQRFVYASSROL 1140
DB 1081 DTWDSMILEVINKPMGSLFNVGACIVTILATPAALIIIPPLIGLYFVQRFVYASSROL 1140

QY 1141 KRLESVRSFVYSHFNFTLLGVSVIRAFEEQERFIHOSDLKVDENQAKYPSIVANRWLA 1200
DB 1141 KRLESVRSFVYSHFNFTLLGVSVIRAFEEQERFIHOSDLKVDENQAKYPSIVANRWLA 1200

QY 1201 VRLCEVNCIVLPALPAVISRHSLSAGLVGSVSYSLQVTTYLNLVWRSSEMETNIVA 1260
DB 1201 VRLCEVNCIVLPALPAVISRHSLSAGLVGSVSYSLQVTTYLNLVWRSSEMETNIVA 1260

QY 1261 VERLKEYSETEKAPMOIOTAPSSWPQVGRVFRNYCLRYREDLDFVLRHINVTINGG 1320
DB 1261 VERLKEYSETEKAPMOIOTAPSSWPQVGRVFRNYCLRYREDLDFVLRHINVTINGG 1320

QY 1321 EKVGIQRTAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
DB 1321 EKVGIQRTAGKSSLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380

QY 1381 SGLSRMNLDPFSGYSDDEVWTSLELAHLKDFVSLPDKLDHECAEGGENLSVGRQVLCL 1440
DB 1381 SGLSRMNLDPFSGYSDDEVWTSLELAHLKDFVSLPDKLDHECAEGGENLSVGRQVLCL 1440

QY 1441 ARALLRKTILVUDEATAVNDLETDDLIOSTIRTOFEDCTVLITAHRLNTIMDTYIVL 1500
DB 1441 ARALLRKTILVUDEATAVNDLETDDLIOSTIRTOFEDCTVLITAHRLNTIMDTYIVL 1500

QY 1501 DKGEIOEYGA PSDLLOOGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGA PSDLLOOGLFYSMAKDAGLV 1531

RESULT 7
US-10-807-466-6
Sequence 6, Application US/10807466
Publication No. US20040244066A1
GENERAL INFORMATION:
APPLICANT: ROTH, CHARLES W.
APPLICANT: BREV, PAUL T.
APPLICANT: HOLM, INGE
APPLICANT: GRALLES, MARINE
APPLICANT: RZHEVSKY, ANDREY
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND
FILE REFERENCE: 03495.0294-01000
CURRENT APPLICATION NUMBER: US/10/807,466
CURRENT FILING DATE: 2004-03-24
PRIOR APPLICATION NUMBER: 10/667,891
PRIOR FILING DATE: 2003-09-23
PRIOR APPLICATION NUMBER: 60/413,469
PRIOR FILING DATE: 2002-09-26
NUMBER OF SEQ ID NOS: 139
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO 6
LENGTH: 1531
TYPE: PRT
ORGANISM: Homo sapiens
US-10-807-466-6

Query Match 80.7%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSDGSDPLMDNMVNTWNTSNPDTKCFONTVLVWPCFYLMACPPFYFLYSRH 60

```
Db 1 MALRGFCSADGSDLDLMDMNTWMTNSNDFTKCFQNTVLLVWPCCYTLMACFPFYLYLSRH 60
Qy 61 DRGVIQMTPLNKTTALGFLIMIVCMADLFYSFWERSRGIFLAIVFLVSPILLGITTLTA 120
Db 61 DRGVIQMTPLNKTTALGFLIMIVCMADLFYSFWERSRGIFLAIVFLVSPILLGITTLTA 120
Qy 121 TFLIQLERRKGVQSSGIMLTFWLVALCALILRSKIMTALKEDAQVDLFPDITFYVYFS 180
Db 121 TFLIQLERRKGVQSSGIMLTFWLVALCALILRSKIMTALKEDAQVDLFPDITFYVYFS 180
Qy 181 LLLIQLVLSGSDSPLESETIHDNPNCPRESSASFLSITFMWITGLIVRGYRQPLBSSD 240
Db 181 LLLIQLVLSGSDSPLESETIHDNPNCPRESSASFLSITFMWITGLIVRGYRQPLBSSD 240
Qy 241 LMSLNKEDTSEQVVPVLYKMKKCAKTRKQPVKVVYSSKXDAOPKSSKYDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLYKMKKCAKTRKQPVKVVYSSKXDAOPKSSKYDANEVEAL 300
Qy 301 IVKSPOKEMNSLKVLYKTGPYFLMSFFPKAIDHLMFSGPOLKLLIFVNDTKAPD 360
Db 301 IVKSPOKEMNSLKVLYKTGPYFLMSFFPKAIDHLMFSGPOLKLLIFVNDTKAPD 360
Qy 361 WQGFYFVTLFVTACLOTVLHOFPHICFVSQMRKTAIVIGAVYRKALVITNSARKSSTV 420
Db 361 WQGFYFVTLFVTACLOTVLHOFPHICFVSQMRKTAIVIGAVYRKALVITNSARKSSTV 420
Qy 421 GEIYVNLNSVDAQRFMDLATYINMTWSAPLQVTLALYLLMLNIGPSVLAVGAVMYLVMPVN 480
Db 421 GEIYVNLNSVDAQRFMDLATYINMTWSAPLQVTLALYLLMLNIGPSVLAVGAVMYLVMPVN 480
Qy 481 AVNANKTKTYVAAHMKSKONRIKLMNEILNGIKVLYKLYAMBLAFKDKYLAIROSELKYLK 540
Db 481 AVNANKTKTYVAAHMKSKONRIKLMNEILNGIKVLYKLYAMBLAFKDKYLAIROSELKYLK 540
Qy 541 KSAVLASVGTFTWCTPPLVALCTPAVYVTTIDENNIIIDAOAPFSLAFNLRPLNLTLP 600
Db 541 KSAVLASVGTFTWCTPPLVALCTPAVYVTTIDENNIIIDAOAPFSLAFNLRPLNLTLP 600
Qy 601 MVISSIVQASVSLKRLRIFLSHEELEPDSERRPVKDGGSITVRNATFTWARSDDPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSERRPVKDGGSITVRNATFTWARSDDPT 660
Qy 661 LMGITFSPREGALVAVVGVQVCGKSSLSALLAEMDKVEGHVAIKGSVAVYVPOQAMIOND 720
Db 661 LMGITFSPREGALVAVVGVQVCGKSSLSALLAEMDKVEGHVAIKGSVAVYVPOQAMIOND 720
Qy 721 SLRENIILFGCOLERPYSRVIOACALLPDLILBPGDTEIGEGVNLISGQOKRVSILAR 780
Db 721 SLRENIILFGCOLERPYSRVIOACALLPDLILBPGDTEIGEGVNLISGQOKRVSILAR 780
Qy 781 AVYSNADIYLFDDPLSAVDAHVKHIFENVIQPKGMLKNKTRILVTHSMSYLPOVDVITV 840
Db 781 AVYSNADIYLFDDPLSAVDAHVKHIFENVIQPKGMLKNKTRILVTHSMSYLPOVDVITV 840
Qy 841 MSGGKISEMGSYQEBLLARDGAPAEFLRTYASTEDQDAEENGVTGVSQPKAKOMENG 900
Db 841 MSGGKISEMGSYQEBLLARDGAPAEFLRTYASTEDQDAEENGVTGVSQPKAKOMENG 900
Qy 901 LVTHSAGKQIOLRSSSSSSSGDTSRHNSSTAELOKAKAKKEBTMKMEADKAOQGVYKL 960
Db 901 LVTHSAGKQIOLRSSSSSSSGDTSRHNSSTAELOKAKAKKEBTMKMEADKAOQGVYKL 960
Qy 961 SVYVDYMKAIQLFISFLSIFLPMCHVSAALASNYWLSMTDDPIVNGIQEHTKRLSYG 1020
Db 961 SVYVDYMKAIQLFISFLSIFLPMCHVSAALASNYWLSMTDDPIVNGIQEHTKRLSYG 1020
Qy 1021 ALGISOGIAVFGYSMAVSIIGILASRCLAHVLDLHSILKSPMSFFERTSPGNLVNRSKEL 1080
Db 1021 ALGISOGIAVFGYSMAVSIIGILASRCLAHVLDLHSILKSPMSFFERTSPGNLVNRSKEL 1080
Qy 1081 DTUDSMIPBVIKMMGSLFNVIAGCTVILATPIAIIIPPLGLIYFVQRYVASSQOL 1140
Db 1081 DTUDSMIPBVIKMMGSLFNVIAGCTVILATPIAIIIPPLGLIYFVQRYVASSQOL 1140
```

```
Qy 1141 KRLESVRSRPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQAVYPSIVANRWLA 1200
Db 1141 KRLESVRSRPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQAVYPSIVANRWLA 1200
Qy 1201 VRLCEVNCIVLPAALPAVVISRHSLSAGLVGLSVSYLOVTTYINMLVRMSSEMETNIVA 1260
Db 1201 VRLCEVNCIVLPAALPAVVISRHSLSAGLVGLSVSYLOVTTYINMLVRMSSEMETNIVA 1260
Qy 1261 VERLKEVSETEKEAPMOIOETAPPSSWPOYGRVEFRNYCLRYREDDLFPVLRHINVTNGG 1320
Db 1261 VERLKEVSETEKEAPMOIOETAPPSSWPOYGRVEFRNYCLRYREDDLFPVLRHINVTNGG 1320
Qy 1321 EKVGIVRTGAGKSSLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKITTIPQDPVLF 1380
Db 1321 EKVGIVRTGAGKSSLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKITTIPQDPVLF 1380
Qy 1381 SSGIRNMLDPPSOVSDEEWMTSLFLAHLKQFVSALPKLDBECAGGBENLSVGQROLVCL 1440
Db 1381 SSGIRNMLDPPSOVSDEEWMTSLFLAHLKQFVSALPKLDBECAGGBENLSVGQROLVCL 1440
Qy 1441 ARALLRKTILVDEAFAAVDLETFDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRYVL 1500
Db 1441 ARALLRKTILVDEAFAAVDLETFDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRYVL 1500
Qy 1501 DKGEIOEYGAPSDLLQQRGLFYSMAKADAGLV 1531
Db 1501 DKGEIOEYGAPSDLLQQRGLFYSMAKADAGLV 1531

RESULT 8
US-10-484-577-678
; Sequence 678, Application US/10484577
; Publication No. US20050032724A1
; GENERAL INFORMATION:
; APPLICANT: EPIDIAKROS Biotechnologie Aktiengesellschaft
; TITLE OF INVENTION: Means and methods for improved treatment of cancer based on UCT1A1
; FILE REFERENCE: F2285PCT-1
; CURRENT APPLICATION NUMBER: US/10/484,577
; CURRENT FILING DATE: 2004-01-22
; PRIOR APPLICATION NUMBER: PCT/EP 02/08220
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: EP 01 11 7608.8
; PRIOR FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: EP 02011710.7
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 683
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 678
; LENGTH: 1531
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-484-577-678

Query Match 80.7%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 241 LMSLNKEDTSEOVVPLVVKWKKCECAKTRKOPVKVYVSSKDPAPKSSKVDANEVEAL 300
Db 241 LMSLNKEDTSEOVVPLVVKWKKCECAKTRKOPVKVYVSSKDPAPKSSKVDANEVEAL 300
QY 301 IVKSPQKEMNPSLFKVLKTFGPFYFLMSFFPKAIHDLMMFGSQPQILKLLIKFVNDTKAPD 360
Db 301 IVKSPQKEMNPSLFKVLKTFGPFYFLMSFFPKAIHDLMMFGSQPQILKLLIKFVNDTKAPD 360
QY 361 WQGFYTYLLFVYLAQTLVLHQYFPHICFVSGMRKTAIVIGAVRKALVTNSARKSSTV 420
Db 361 WQGFYTYLLFVYLAQTLVLHQYFPHICFVSGMRKTAIVIGAVRKALVTNSARKSSTV 420
QY 421 GEIVNLMSVDAORFMDLATTYINMTMSAPLOVTLAITYLMLNGSPVLAGVAVMTLMVAVN 480
Db 421 GEIVNLMSVDAORFMDLATTYINMTMSAPLOVTLAITYLMLNGSPVLAGVAVMTLMVAVN 480
QY 481 AVNMMKTKTYOVAAHMSKDNRIKLMNEILNGIKVLKLYAMELAFDKYLAIROBELKYLK 540
Db 481 AVNMMKTKTYOVAAHMSKDNRIKLMNEILNGIKVLKLYAMELAFDKYLAIROBELKYLK 540
QY 541 KSAVLSAVGTFTWCTPPLVALCTPAVYVTIDENNILDAQTAFYSLALFNILRPDLNLP 600
Db 541 KSAVLSAVGTFTWCTPPLVALCTPAVYVTIDENNILDAQTAFYSLALFNILRPDLNLP 600
QY 601 MVISIVQASVSLKRLRIEFLSHEELPDISERRPVKDGCGNSTTVRAATTTWASDPT 660
Db 601 MVISIVQASVSLKRLRIEFLSHEELPDISERRPVKDGCGNSTTVRAATTTWASDPT 660
QY 661 LMGITFSPBEGALVAVVGVCCKSSLSALLAEMDKYGHVAIKGSVAVYPOQAMIOND 720
Db 661 LMGITFSPBEGALVAVVGVCCKSSLSALLAEMDKYGHVAIKGSVAVYPOQAMIOND 720
QY 721 SLRENILFGCOLLEBPYRSVIOACALLPDLILPBGDTEIGEGVNLSSGQOKRVSLAR 780
Db 721 SLRENILFGCOLLEBPYRSVIOACALLPDLILPBGDTEIGEGVNLSSGQOKRVSLAR 780
QY 781 AVYNNADITYLEPDDPLSANDAHVGHIFENVIQPKMKNKRIIVLTHSMSTLPQVDVITV 840
Db 781 AVYNNADITYLEPDDPLSANDAHVGHIFENVIQPKMKNKRIIVLTHSMSTLPQVDVITV 840
QY 841 MSGGKISEMGSYOELLARDGAFAEFLRTYASTEOBDAEENGVTVSGPGKEAKOMENGM 900
Db 841 MSGGKISEMGSYOELLARDGAFAEFLRTYASTEOBDAEENGVTVSGPGKEAKOMENGM 900
QY 901 LVTDSAGKOLQROLSSSSSYSGDISRHNSSTAELOKAEKKEETWKLMEADKAOTGVKL 960
Db 901 LVTDSAGKOLQROLSSSSSYSGDISRHNSSTAELOKAEKKEETWKLMEADKAOTGVKL 960
QY 961 SYVYNDYMKALIGLFSLSIFLPMCNHVSALASNTWLSMTDDPIVNGTOETKRLSYG 1020
Db 961 SYVYNDYMKALIGLFSLSIFLPMCNHVSALASNTWLSMTDDPIVNGTOETKRLSYG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLHVDLHSLIRSPMSFFERPSPGULVNRFSKEL 1080
Db 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLHVDLHSLIRSPMSFFERPSPGULVNRFSKEL 1080
QY 1081 DTVDSMTPEVILKMGSLFENVIGACIVILLATPIAIIIPPLGLIYEFVORFYVASSROL 1140
Db 1081 DTVDSMTPEVILKMGSLFENVIGACIVILLATPIAIIIPPLGLIYEFVORFYVASSROL 1140
QY 1141 KRLESVRSRPSYSHFNETLLGVSVITRAFEBOERFIHOSDLKVDENOKAYYSIYANRWLA 1200
Db 1141 KRLESVRSRPSYSHFNETLLGVSVITRAFEBOERFIHOSDLKVDENOKAYYSIYANRWLA 1200
QY 1201 VRLECVGNCTVLFALFAVISRHSLSAGLVGLSVYSLOVTTYLNMILVRMSSMETNIVA 1260
Db 1201 VRLECVGNCTVLFALFAVISRHSLSAGLVGLSVYSLOVTTYLNMILVRMSSMETNIVA 1260
QY 1261 VERLKEYSETEKAPMOJQETAPSSWPQVGRVEFRNYCLTRYREDLDFVLRIHINVTTNGG 1320
Db 1261 VERLKEYSETEKAPMOJQETAPSSWPQVGRVEFRNYCLTRYREDLDFVLRIHINVTTNGG 1320

QY 1321 EKVGIIVGTAGKSSLTGLFRINESAGEIILIDGINAKIGLHDLRFKTIIPODPVLF 1380
Db 1321 EKVGIIVGTAGKSSLTGLFRINESAGEIILIDGINAKIGLHDLRFKTIIPODPVLF 1380
QY 1381 SGSLRMLNDPFSQYSDEEWMTSLBLAHKDFVSALPDQLDHECAEGENLSVGROLYCL 1440
Db 1381 SGSLRMLNDPFSQYSDEEWMTSLBLAHKDFVSALPDQLDHECAEGENLSVGROLYCL 1440
QY 1441 ARALLRKTILVDEBATAVDLETDLIQSTIRIQFEDCTVTLTAHRLNTIMDTYRIVL 1500
Db 1441 ARALLRKTILVDEBATAVDLETDLIQSTIRIQFEDCTVTLTAHRLNTIMDTYRIVL 1500
QY 1501 DKGEIOEGAPSDLLQORGLFYSMAXDAGV 1531
Db 1501 DKGEIOEGAPSDLLQORGLFYSMAXDAGV 1531

RESULT 9
US-10-889-503-19
; Sequence 19, Application US/10889503
; Publication No. US20050063968A1
; GENERAL INFORMATION:
; APPLICANT: Fox Chase Cancer Center
; APPLICANT: Krush, Gary D.
; APPLICANT: Lee, Kun
; APPLICANT: Belinsky, Martin G.
; TITLE OF INVENTION: MRP-Related ABC Transporter Encoding
; FILE REFERENCE: FCCC 98-02
; CURRENT APPLICATION NUMBER: US/10/889, 503
; PRIOR FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: US/09/647, 140
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US99/06644
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/079, 759
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/095, 153
; PRIOR FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-889-503-19

Query Match 80.7%; Score 7860; DB 5; Length 1531;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCAGSDPLDMNNTANTSNPDKCFONTLVWVPCYLMACPPFYLYSRH 60
Db 1 MALRGFCAGSDPLDMNNTANTSNPDKCFONTLVWVPCYLMACPPFYLYSRH 60
QY 61 DRGYIQMTPLNKTKTALGFLIMIVCMADLFYFWEERSRGIFLAEVPLVSPTLGITTLA 120
Db 61 DRGYIQMTPLNKTKTALGFLIMIVCMADLFYFWEERSRGIFLAEVPLVSPTLGITTLA 120
QY 121 TPLIQLERRKGVQSSGIMLTFMVALVCAALILRSKIMTLAKEDAOVDLFRDITFYVYS 180
Db 121 TPLIQLERRKGVQSSGIMLTFMVALVCAALILRSKIMTLAKEDAOVDLFRDITFYVYS 180
QY 181 LLLILQVLSCFSPBSPLFSEITIHDPNCPRESSASFLSRTIFMMITGLIVGYROPLBSPD 240
Db 181 LLLILQVLSCFSPBSPLFSEITIHDPNCPRESSASFLSRTIFMMITGLIVGYROPLBSPD 240
QY 241 LMSLNKEDTSEOVVPLVVKWKKCECAKTRKOPVKVYVSSKDPAPKSSKVDANEVEAL 300
Db 241 LMSLNKEDTSEOVVPLVVKWKKCECAKTRKOPVKVYVSSKDPAPKSSKVDANEVEAL 300
QY 301 IVKSPQKEMNPSLFKVLKTFGPFYFLMSFFPKAIHDLMMFGSQPQILKLLIKFVNDTKAPD 360

```
Db 301 IVKSPQKEMNSLFFVLYKTEGPFYLMSPFKAIHDLMMFSGPOLIKLIRFVNDTKAPD 360
Qy 361 WQGFYTYVLLFVYACLOTLVYHOYFHIQFVSGMRIKTAVIGAVYKALVITNSARKSSTV 420
Db 361 WQGFYTYVLLFVYACLOTLVYHOYFHIQFVSGMRIKTAVIGAVYKALVITNSARKSSTV 420
Qy 421 GEIVNLSVDAQRFMDLATYINMISAPLOYLALYLLMLNGPSVLGAVAVMLMVEVN 480
Db 421 GEIVNLSVDAQRFMDLATYINMISAPLOYLALYLLMLNGPSVLGAVAVMLMVEVN 480
Qy 481 AVMAKTKTYQVAHMKSDNR1KLMNEILNGIKYVLAUYAMELAFKDYLAIRQEBLKYLK 540
Db 481 AVMAKTKTYQVAHMKSDNR1KLMNEILNGIKYVLAUYAMELAFKDYLAIRQEBLKYLK 540
Qy 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVTDENNILDAQAFVSLFENILRPLNTLP 600
Db 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVTDENNILDAQAFVSLFENILRPLNTLP 600
Qy 601 MVISIVQASVSLKRLRIFLSHELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
Db 601 MVISIVQASVSLKRLRIFLSHELEPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
Qy 661 LMGITFSIPGAGLVAVVQVQGGKSSLSALLAEMDKYEGHVAIGSVAYVPOQAMIOND 720
Db 661 LMGITFSIPGAGLVAVVQVQGGKSSLSALLAEMDKYEGHVAIGSVAYVPOQAMIOND 720
Qy 721 SLRENIIFGCOLBEEYYRSVIOACALPDLLEILPSGDRTEIOEKGVNLSGGOKOVSILAR 780
Db 721 SLRENIIFGCOLBEEYYRSVIOACALPDLLEILPSGDRTEIOEKGVNLSGGOKOVSILAR 780
Qy 781 AVYSNADYILFPDDLPAVDAAHVGKHI FENVIGPKMLNKNKRLIVTSHMSYLPQVDVIIV 840
Db 781 AVYSNADYILFPDDLPAVDAAHVGKHI FENVIGPKMLNKNKRLIVTSHMSYLPQVDVIIV 840
Qy 841 MSGKISMGSTYOELLARDGAFBLRTYASTEOQDAEENGVTGVSQGGKAKOMENGM 900
Db 841 MSGKISMGSTYOELLARDGAFBLRTYASTEOQDAEENGVTGVSQGGKAKOMENGM 900
Qy 901 LVTDSAGOLOROSSSSSGDISRHHNSTAELOKAEKKEETKMLWEADAAQOYVYL 960
Db 901 LVTDSAGOLOROSSSSSGDISRHHNSTAELOKAEKKEETKMLWEADAAQOYVYL 960
Qy 961 SVYMDYMKAIGLFISFLSIFLFMCNHSALASNYMLSLMTDPIVNGTOEHTKVALSYVG 1020
Db 961 SVYMDYMKAIGLFISFLSIFLFMCNHSALASNYMLSLMTDPIVNGTOEHTKVALSYVG 1020
Qy 1021 ALGISQGIADVGYSAVSIIGTILASRCLHVDLHSLNSPMSFFERTPSGNLVNRFSEKL 1080
Db 1021 ALGISQGIADVGYSAVSIIGTILASRCLHVDLHSLNSPMSFFERTPSGNLVNRFSEKL 1080
Qy 1081 DTVDSMIEPVIKMFNGSLFENVIGACIVILATPRAIITIPGLIYFFQRFYVASSRQL 1140
Db 1081 DTVDSMIEPVIKMFNGSLFENVIGACIVILATPRAIITIPGLIYFFQRFYVASSRQL 1140
Qy 1141 KRLBSVSHSPVYSHNETLIGSVIRAFEOBRFTHQSDLKVDENOKAYPSIVANRWLA 1200
Db 1141 KRLBSVSHSPVYSHNETLIGSVIRAFEOBRFTHQSDLKVDENOKAYPSIVANRWLA 1200
Qy 1201 VRLBECVNCIVLFAALFAVISHSLSAGLVGLSVSYSLQVTTYLMLVMSSEMETNIVA 1260
Db 1201 VRLBECVNCIVLFAALFAVISHSLSAGLVGLSVSYSLQVTTYLMLVMSSEMETNIVA 1260
Qy 1261 VERLKEVETEKEAWOIOETAPRSSWPOVGRVPRPNYCLARYREDLDFVLRHINVTING 1320
Db 1261 VERLKEVETEKEAWOIOETAPRSSWPOVGRVPRPNYCLARYREDLDFVLRHINVTING 1320
Qy 1321 EKVGIVGRTAGKSSULTGLFRINSABEIIIDGINAKIGLHDLRFKITIIPODPVLF 1380
Db 1321 EKVGIVGRTAGKSSULTGLFRINSABEIIIDGINAKIGLHDLRFKITIIPODPVLF 1380
Qy 1381 SGSLMNI.DPFSQYSDEBVTSLLELAHLKD FVSALPDKLDHECABEGENLSVGORQVCL 1440
Db 1381 SGSLMNI.DPFSQYSDEBVTSLLELAHLKD FVSALPDKLDHECABEGENLSVGORQVCL 1440
```

```
Db 1381 SGSLMNI.DPFSQYSDEBVTSLLELAHLKD FVSALPDKLDHECABEGENLSVGORQVCL 1440
Qy 1441 ARALIRKTKILVUDEATAVADLETDLIQSTIRFOEDCTVLIARHLNTIMDYTRVIL 1500
Db 1441 ARALIRKTKILVUDEATAVADLETDLIQSTIRFOEDCTVLIARHLNTIMDYTRVIL 1500
Qy 1501 DKGEIOBYGAPSDLLQORGLFYMAKODAGLV 1531
Db 1501 DKGEIOBYGAPSDLLQORGLFYMAKODAGLV 1531

RESULT 10
US-10-756-149-5033
; Sequence 5033, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Naeem
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756.149
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5033
; LENGTH: 1531
; TYPE: PRF
; ORGANISM: Homo Sapiens
US-10-756-149-5033

Query Match 80.7%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRFGCADSDPLMDMNTWNTSNPFTYCFQNTYLVWVPCRYLWACFPFFLYLSRH 60
Db 1 MALRFGCADSDSDPLMDMNTWNTSNPFTYCFQNTYLVWVPCRYLWACFPFFLYLSRH 60
Qy 61 DRGYIOWMPLNKTKTALGFLIMYCMADLFYSPMERSGIFLAPVFLVSPILGITTLLA 120
Db 61 DRGYIOWMPLNKTKTALGFLIMYCMADLFYSPMERSGIFLAPVFLVSPILGITTLLA 120
Qy 121 TFLIQLERRKVGQSGIMLTFMVLVYCALAIIKSKIMTALKEDAQVDLPFDITFYVYS 180
Db 121 TFLIQLERRKVGQSGIMLTFMVLVYCALAIIKSKIMTALKEDAQVDLPFDITFYVYS 180
Qy 181 LLLIOLVLSCTSDNSPLFSETIHDNPPCPRESSASFLSKITTWMTTGLIVRGYRQPLEGSD 240
Db 181 LLLIOLVLSCTSDNSPLFSETIHDNPPCPRESSASFLSKITTWMTTGLIVRGYRQPLEGSD 240
Qy 241 LMSLNKEDTSEOVVAVLVKMKKECAKTRKOPVKVYSSXDPAPQKSSKYDANEVBAL 300
Db 241 LMSLNKEDTSEOVVAVLVKMKKECAKTRKOPVKVYSSXDPAPQKSSKYDANEVBAL 300
Qy 301 IVKSPQKEMNSLFFVLYKTEGPFYLMSPFKAIHDLMMFSGPOLIKLIRFVNDTKAPD 360
Db 301 IVKSPQKEMNSLFFVLYKTEGPFYLMSPFKAIHDLMMFSGPOLIKLIRFVNDTKAPD 360
Qy 361 WQGFYTYVLLFVYACLOTLVYHOYFHIQFVSGMRIKTAVIGAVYKALVITNSARKSSTV 420
Db 361 WQGFYTYVLLFVYACLOTLVYHOYFHIQFVSGMRIKTAVIGAVYKALVITNSARKSSTV 420
Qy 421 GEIVNLSVDAQRFMDLATYINMISAPLOYLALYLLMLNGPSVLGAVAVMLMVEVN 480
Db 421 GEIVNLSVDAQRFMDLATYINMISAPLOYLALYLLMLNGPSVLGAVAVMLMVEVN 480
Qy 481 AVMAKTKTYQVAHMKSDNR1KLMNEILNGIKYVLAUYAMELAFKDYLAIRQEBLKYLK 540
Db 481 AVMAKTKTYQVAHMKSDNR1KLMNEILNGIKYVLAUYAMELAFKDYLAIRQEBLKYLK 540
Qy 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVTDENNILDAQAFVSLFENILRPLNTLP 600
Db 541 KSAVLSAVGTFTWCTPPLVALCTFAVYVTDENNILDAQAFVSLFENILRPLNTLP 600
```

Db 541 KSAVLSAGFTWVCTPFLVALCTFAVYVTIDENNILDAQTAFAVSLAFNIRPLNLP 600
 Qy 601 MVSIIYOASVSLRLRIPLSHEELPDSIERRPVKGGGNSITVNAATTTARSPPT 660
 Db 601 MVSIIYOASVSLRLRIPLSHEELPDSIERRPVKGGGNSITVNAATTTARSPPT 660
 Qy 661 LNTGTFPIPSGALVAVVGVCGGKSSLSALLAEMDKVGHVAKGSVAVYPOQAMQND 720
 Db 661 LNTGTFPIPSGALVAVVGVCGGKSSLSALLAEMDKVGHVAKGSVAVYPOQAMQND 720
 Qy 721 SLRENIIIFGCOLPEPPYRSVIAQACALLPDLIELPSGDRTEIGEGVNLSSGGQKRVSLAR 780
 Db 721 SLRENIIIFGCOLPEPPYRSVIAQACALLPDLIELPSGDRTEIGEGVNLSSGGQKRVSLAR 780
 Qy 781 AVYSNADIIYLPDDLSAVDAVGHRIPENVIGPGKMLKNKRIILVTHSMSTLPQVDYIIV 840
 Db 781 AVYSNADIIYLPDDLSAVDAVGHRIPENVIGPGKMLKNKRIILVTHSMSTLPQVDYIIV 840
 Qy 841 MSGGKISEMGSYOELLARDGAFAEFLRTYASTEOEODAEENGVTGVSQPKAKQOMENG 900
 Db 841 MSGGKISEMGSYOELLARDGAFAEFLRTYASTEOEODAEENGVTGVSQPKAKQOMENG 900
 Qy 901 LYTDSACKOLQROQLSSSSSYSGDISRHHNSTAELQKAKKEETWKLMEADKAQTGVKL 960
 Db 901 LYTDSACKOLQROQLSSSSSYSGDISRHHNSTAELQKAKKEETWKLMEADKAQTGVKL 960
 Qy 961 SYVWDYWKALGLFISFLISFLFMCNHSALASNTWLSMTDDPIVNGTOEHTKRLSVYG 1020
 Db 961 SYVWDYWKALGLFISFLISFLFMCNHSALASNTWLSMTDDPIVNGTOEHTKRLSVYG 1020
 Qy 1021 ALGISOGIAVFGYSMAVISGIIASRCIAHYDLHSIIRSPMSPEERTPSGKLVNRFSEKL 1080
 Db 1021 ALGISOGIAVFGYSMAVISGIIASRCIAHYDLHSIIRSPMSPEERTPSGKLVNRFSEKL 1080
 Qy 1081 DTVDMSIPEVYIKMGMSLFNVIGACIYIILATPIAIIIPPLGLIYFVGQRYVASSROL 1140
 Db 1081 DTVDMSIPEVYIKMGMSLFNVIGACIYIILATPIAIIIPPLGLIYFVGQRYVASSROL 1140
 Qy 1141 KRLSVRSRPSVYSHFNFTLLGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANMFLA 1200
 Db 1141 KRLSVRSRPSVYSHFNFTLLGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANMFLA 1200
 Qy 1201 VRLKESVRSRPSVYSHFNFTLLGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANMFLA 1260
 Db 1201 VRLKESVRSRPSVYSHFNFTLLGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANMFLA 1260
 Qy 1261 VERLKEYSETEKAPMOIOTETAPSSWPQYGRVFRNYCLRYEDDLFVLRIHNVITNGG 1320
 Db 1261 VERLKEYSETEKAPMOIOTETAPSSWPQYGRVFRNYCLRYEDDLFVLRIHNVITNGG 1320
 Qy 1321 EKVGIYGRTAGKSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
 Db 1321 EKVGIYGRTAGKSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
 Qy 1381 SSGSLRMLDPSOYSDEEWTSLIELAKDFVSALPKLDHECAGEGENSVGQROVLCL 1440
 Db 1381 SSGSLRMLDPSOYSDEEWTSLIELAKDFVSALPKLDHECAGEGENSVGQROVLCL 1440
 Qy 1441 ARALLKRTKILVDEATAVDELTDDLIQSTIRTOFEDCTVLTIAHLANTIMOTRYIVL 1500
 Db 1441 ARALLKRTKILVDEATAVDELTDDLIQSTIRTOFEDCTVLTIAHLANTIMOTRYIVL 1500
 Qy 1501 DKGEIOEGAPSDLLQORGLFYSMADAGLIV 1531
 Db 1501 DKGEIOEGAPSDLLQORGLFYSMADAGLIV 1531

; TITLE OF INVENTION: No. US20040039163a1el Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-099
 ; CURRENT APPLICATION NUMBER: US/09/939,853A
 ; PRIOR FILING DATE: 2001-08-27
 ; PRIOR APPLICATION NUMBER: 60/228,191
 ; PRIOR FILING DATE: 2000-08-25
 ; PRIOR APPLICATION NUMBER: 60/267,300
 ; PRIOR FILING DATE: 2001-02-08
 ; PRIOR APPLICATION NUMBER: 60/269,961
 ; PRIOR FILING DATE: 2001-02-20
 ; PRIOR APPLICATION NUMBER: 60/277,337
 ; PRIOR FILING DATE: 2001-03-20
 ; NUMBER OF SEQ ID NOS: 159
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 87
 ; LENGTH: 1515
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-939-853A-87
 Query Match 79.8%; Score 7769; DB 3; Length 1515;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 17 DMVNTWNTSMPDFTKCFQNTVLVWVPCFYLMAQCPFFYFLSHRDREYIOMTPINKTKTA 76
 Db 1 DMVNTWNTSMPDFTKCFQNTVLVWVPCFYLMAQCPFFYFLSHRDREYIOMTPINKTKTA 60
 Qy 77 LGFLMLVCAADLFYFSEWERSRGIPLAPVPLVSGPTLLGTTLLATFLIQERRKGVOSG 136
 Db 61 LGFLMLVCAADLFYFSEWERSRGIPLAPVPLVSGPTLLGTTLLATFLIQERRKGVOSG 120
 Qy 137 IMLTFWLVALVCAIIRSKIMTALKEADAQVDLFRDITFYVYVSLLIQLVSCFSRSP 196
 Db 121 IMLTFWLVALVCAIIRSKIMTALKEADAQVDLFRDITFYVYVSLLIQLVSCFSRSP 180
 Qy 197 LFSETHDPNCPRESSASFLSRTTFWWTGLVYRGROPLSGDLSLNEEDTSQVVPV 256
 Db 181 LFSETHDPNCPRESSASFLSRTTFWWTGLVYRGROPLSGDLSLNEEDTSQVVPV 240
 Qy 257 LVNKMKECKATKQKOPKVVYSSKDPAPQKSSKVDANEVEALIVKSPKEMNPSLFKV 316
 Db 241 LVNKMKECKATKQKOPKVVYSSKDPAPQKSSKVDANEVEALIVKSPKEMNPSLFKV 300
 Qy 317 LYTFPGPYFLMSEFFRAIHDLMFSGPQIILKLIKPVNDTKADWQGYFTVLLFVTACL 376
 Db 301 LYTFPGPYFLMSEFFRAIHDLMFSGPQIILKLIKPVNDTKADWQGYFTVLLFVTACL 360
 Qy 377 QTLVLAQYFHICVSGMRITANTAVIGAVYRQALVITNSARKSSVGEIVNLSVDAQRFMD 436
 Db 361 QTLVLAQYFHICVSGMRITANTAVIGAVYRQALVITNSARKSSVGEIVNLSVDAQRFMD 420
 Qy 437 LATYINMINSAPQVILALYILMLNGPSVLAVAVVWLVPNVAWAMTKTYQVVAHMK 496
 Db 421 LATYINMINSAPQVILALYILMLNGPSVLAVAVVWLVPNVAWAMTKTYQVVAHMK 480
 Qy 497 SKDNRIKLAMEILNGIKVLKLYAMELAFKDKVLAIRQEBELKVLKKSAYLSAVGFTWVCT 556
 Db 481 SKDNRIKLAMEILNGIKVLKLYAMELAFKDKVLAIRQEBELKVLKKSAYLSAVGFTWVCT 540
 Qy 557 PFLVALCTPRAVYTTIDENNILDAQTAFAVSLAFNIRPLNLPMTYISIVQASVSLKRL 616
 Db 541 PFLVALCTPRAVYTTIDENNILDAQTAFAVSLAFNIRPLNLPMTYISIVQASVSLKRL 600
 Qy 617 RIFLSHEELPDSIERRPVKGGGNSITVNAATTTARSPPTLNGITSPISGALVAV 676
 Db 601 RIFLSHEELPDSIERRPVKGGGNSITVNAATTTARSPPTLNGITSPISGALVAV 660
 Qy 677 VGVGCGKSSLSALLAEMDKVGHVAKGSVAVYPOQAMQNDLSRENILFCQOLEEPPY 736
 Db 661 VGVGCGKSSLSALLAEMDKVGHVAKGSVAVYPOQAMQNDLSRENILFCQOLEEPPY 720
 Qy 737 YRSVIAQACALLPDLIELPSGDRTEIGEGVNLSSGGQKRVSLARAVYSNADIIYLPDDPLS 796

RESULT 11
 US-09-939-853A-87
 ; Sequence 87, Application US/09939853A
 ; Publication No. US20040039163a1
 ; GENERAL INFORMATION:
 ; APPLICANT: Burgees et al.

```
|||||
Db 721 YRSVIOACALLPDLIELIPSGDRTIEIGEGVNLGGQKORVSLARAVYSNADIYFDPLSL 780
QY 797 AVDAHVGKHIFENYIGPKGMLKNKTRILIVTHSMSTLPQVDYIIVMSGGKISEMSYQEL 856
Db 781 AVDAHVGKHIFENYIGPKGMLKNKTRILIVTHSMSTLPQVDYIIVMSGGKISEMSYQEL 840
QY 857 ARDGAFAELFRTYASTBEOQDAEENGVTGSGPGKEAKOMENGMVLTDSAGKQLOROLSS 916
Db 841 ARDGAFAELFRTYASTBEOQDAEENGVTGSGPGKEAKOMENGMVLTDSAGKQLOROLSS 900
QY 917 SSSYSGDISRRHNSTAEIQKAEAKKEETWKLMADKATQGVKLSVYVDYKAIIGLFISSF 976
Db 901 SSSYSGDISRRHNSTAEIQKAEAKKEETWKLMADKATQGVKLSVYVDYKAIIGLFISSF 960
QY 977 LSIPLFMCHNYSALASNYWMLSLMTDDPIVNGTOETHKRLSVYGAIGISOGIAVFGYSMA 1036
Db 961 LSIPLFMCHNYSALASNYWMLSLMTDDPIVNGTOETHKRLSVYGAIGISOGIAVFGYSMA 1020
QY 1037 VSIQGIILASRCLAHVDLHSLIRSPMSFPERTPSGNLVNRFSEKLDPTVDSMIPEVIXMFMG 1096
Db 1021 VSIQGIILASRCLAHVDLHSLIRSPMSFPERTPSGNLVNRFSEKLDPTVDSMIPEVIXMFMG 1080
QY 1097 SLENVIGACIYIILATPIAATIIIPPLGIYFPVQAFYVASSRQLKRLSVSRSPYSHFN 1156
Db 1081 SLENVIGACIYIILATPIAATIIIPPLGIYFPVQAFYVASSRQLKRLSVSRSPYSHFN 1140
QY 1157 ETLGVSIVTRAPEEOERFIHOSDLKVDENOKAYYPSIVANRWLAVRLECVGNCIYLPAL 1216
Db 1141 ETLGVSIVTRAPEEOERFIHOSDLKVDENOKAYYPSIVANRWLAVRLECVGNCIYLPAL 1200
QY 1217 PAVISRHSLSAGVLGVLSVYSLQVTTYLNMLVRMSSEMETNIIVAEVLKEYSETEKAPW 1276
Db 1201 PAVISRHSLSAGVLGVLSVYSLQVTTYLNMLVRMSSEMETNIIVAEVLKEYSETEKAPW 1260
QY 1277 QIOETAPSSMPQVGRVFRFNYCLARYRDLDPVLRHINVTINGEKVGIYGRGTGAKSSL 1336
Db 1261 QIOETAPSSMPQVGRVFRFNYCLARYRDLDPVLRHINVTINGEKVGIYGRGTGAKSSL 1320
QY 1337 TLGFRIINESAGEIIGININAKIGLHDLRPKTIIPQDPVLFSGSLRMLNDPPSOYSD 1396
Db 1321 TLGFRIINESAGEIIGININAKIGLHDLRPKTIIPQDPVLFSGSLRMLNDPPSOYSD 1380
QY 1397 BEVWTSLELAHKDFVSALPKLDHECAGEGENTSVGQROLVCLARALLRKTKILVLDEA 1456
Db 1381 BEVWTSLELAHKDFVSALPKLDHECAGEGENTSVGQROLVCLARALLRKTKILVLDEA 1440
QY 1457 TAAVDLETDLIGSTIRTOFEDCTVLTIAHRLNTIMDYTRYIVLDKGEIOEYGAPSDLLQ 1516
Db 1441 TAAVDLETDLIGSTIRTOFEDCTVLTIAHRLNTIMDYTRYIVLDKGEIOEYGAPSDLLQ 1500
QY 1517 ORGLFYSMAXDAGLV 1531
Db 1501 ORGLFYSMAXDAGLV 1515

RESULT 12
US-10-618-281-42
; Sequence 42, Application US/10618281
; Publication No. US20040219609A1
; GENERAL INFORMATION:
; APPLICANT: Day, Anthony G.
; APPLICANT: Estell, David A.
; APPLICANT: Lyons, Eric H.
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: Methods for Modulating Proteins Not
; TITLE OF INVENTION: Previously Known as Proteases
; FILE REFERENCE: GC773-2
; CURRENT APPLICATION NUMBER: US/10/618,281
; PRIOR FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: US 60/395,325
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 92
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 1515
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-618-281-42

Query Match          79.8%; Score 7763; DB 5; Length 1515;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1514; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 17 DNANTNTSPDPFKCFQNTLVVWVPCFYLMACPFYFLYSRRDRGYIOWTPLNKTKTA 76
Db 1 DNVNTWNTSNDPFKCFQNTLVVWVPCFYLMACPFYFLYSRRDRGYIOWTPLNKTKTA 60
QY 77 LGFLMTIVCMADLEFYSWERSRGIFLAPVFLVSPFLGIIITLTLATFLIQLERRKGVSSG 136
Db 61 LGFLMTIVCMADLEFYSWERSRGIFLAPVFLVSPFLGIIITLTLATFLIQLERRKGVSSG 120
QY 137 IMLFMWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYFSLLLIQVLSCFSDRSP 196
Db 121 IMLFMWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYFSLLLIQVLSCFSDRSP 180
QY 197 LFSETIHDNPPCESSSASFLSRITFWMITGLIVRGYQOPLEGSDLMSLNKEDTSEQVVPV 256
Db 181 LFSETIHDNPPCESSSASFLSRITFWMITGLIVRGYQOPLEGSDLMSLNKEDTSEQVVPV 240
QY 257 LVKMKKECAKTRKQOPKVVYSSKDPAQPKSSKYDANEEVEALIVKSPQKEMNPSLFKV 316
Db 241 LVKMKKECAKTRKQOPKVVYSSKDPAQPKSSKYDANEEVEALIVKSPQKEMNPSLFKV 300
QY 317 LYKTFGPFYFLMSFFPKAIHDLMSGPOILKLIKFVNDTKAPDMQGYFTVLLFTVACL 376
Db 301 LYKTFGPFYFLMSFFPKAIHDLMSGPOILKLIKFVNDTKAPDMQGYFTVLLFTVACL 360
QY 377 QTLVLAQYFHI CFPSGMRIRKTAIVGAYYRKALVYTNARKSTYGEIYVNLMSVDAQRPMD 436
Db 361 QTLVLAQYFHI CFPSGMRIRKTAIVGAYYRKALVYTNARKSTYGEIYVNLMSVDAQRPMD 420
QY 437 LATYNNMWSAPLOVIALYLLMLNLPSTVAGVAVMWLPVNAVAMKTKTYQVAHMK 496
Db 421 LATYNNMWSAPLOVIALYLLMLNLPSTVAGVAVMWLPVNAVAMKTKTYQVAHMK 480
QY 497 SKDNRIKIMNEILINGIKVLKYAMELAFKDVLAIROBELKVLKKSAYLSAVGFTWCT 556
Db 481 SKDNRIKIMNEILINGIKVLKYAMELAFKDVLAIROBELKVLKKSAYLSAVGFTWCT 540
QY 557 PFLVALCTFANYVTIDENNIIDAQTAFVSALFNILRFPNLILPMVTSIYQASVSLKRL 616
Db 541 PFLVALCTFANYVTIDENNIIDAQTAFVSALFNILRFPNLILPMVTSIYQASVSLKRL 600
QY 617 RIFLSHEELBPSIERRPVKDGGSNTSYVNAATFTWARSDPPTLNGITFSIPGALYAV 676
Db 601 RIFLSHEELBPSIERRPVKDGGSNTSYVNAATFTWARSDPPTLNGITFSIPGALYAV 660
QY 677 VGQVCGCKSSLISALLAEMDKVEGHVAKGSVAVVPQAWIQNDSLRENIIFGQLEBPY 736
Db 661 VGQVCGCKSSLISALLAEMDKVEGHVAKGSVAVVPQAWIQNDSLRENIIFGQLEBPY 720
QY 737 YRSVIOACALLPDLIELIPSGDRTIEIGEGVNLGGQKORVSLARAVYSNADIYFDPLSL 796
Db 721 YRSVIOACALLPDLIELIPSGDRTIEIGEGVNLGGQKORVSLARAVYSNADIYFDPLSL 780
QY 797 AVDAHVGKHIFENYIGPKGMLKNKTRILIVTHSMSTLPQVDYIIVMSGGKISEMSYQEL 856
Db 781 AVDAHVGKHIFENYIGPKGMLKNKTRILIVTHSMSTLPQVDYIIVMSGGKISEMSYQEL 840
QY 857 ARDGAFAELFRTYASTBEOQDAEENGVTGSGPGKEAKOMENGMVLTDSAGKQLOROLSS 916
Db 841 ARDGAFAELFRTYASTBEOQDAEENGVTGSGPGKEAKOMENGMVLTDSAGKQLOROLSS 900
QY 917 SSSYSGDISRRHNSTAEIQKAEAKKEETWKLMADKATQGVKLSVYVDYKAIIGLFISSF 976
Db 901 SSSYSGDISRRHNSTAEIQKAEAKKEETWKLMADKATQGVKLSVYVDYKAIIGLFISSF 960
```

Db 901 SSSYSGDISRHHNSTAELOKAEKKEETWKLMEADKQGVKLSVMDYMKAIIGLFIISF 960
Qy 977 LSIPLFMCNHYASALASNYWLSLMTDDPIVNGQOHTKXRLSVYGALGISOGIAVPGYMA 1036
Db 961 LSIPLFMCNHYASALASNYWLSLMTDDPIVNGQOHTKXRLSVYGALGISOGIAVPGYMA 1020
Qy 1037 VSIIGILASRCLHVDLHLSILRSPMSFPERTPSGNLVNRFSKELDTVDSMIPEVIMKMG 1096
Db 1021 VSIIGILASRCLHVDLHLSILRSPMSFPERTPSGNLVNRFSKELDTVDSMIPEVIMKMG 1080
Qy 1097 SLFNVIGACIYILLATPIAIIIPPLGLIYFFVQRFYVASSRQKRLSVSRSPVYSHN 1156
Db 1081 SLFNVIGACIYILLATPIAIIIPPLGLIYFFVQRFYVASSRQKRLSVSRSPVYSHN 1140
Qy 1157 ETLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLAVRLECGNCTVLPAL 1216
Db 1141 ETLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLAVRLECGNCTVLPAL 1200
Qy 1217 FAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRRMSSEMETNIVAVERLKEYSETEKEAPW 1276
Db 1201 FAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRRMSSEMETNIVAVERLKEYSETEKEAPW 1260
Qy 1277 QIQETAPSSWPQVGRVFEFRNYCLRYREDLPVLRIHINVTINGEKVGI VGRGTAGKSSL 1336
Db 1261 QIQETAPSSWPQVGRVFEFRNYCLRYREDLPVLRIHINVTINGEKVGI VGRGTAGKSSL 1320
Qy 1337 TLGLFRINESAEGEIIIDGINIAKIGLHDLAFKTTIIPQDBVLVSSGSRMLNDPSSQYSD 1396
Db 1321 TLGLFRINESAEGEIIIDGINIAKIGLHDLAFKTTIIPQDBVLVSSGSRMLNDPSSQYSD 1380
Qy 1397 EEWTSLELAHLKDFVSALPKDLHECAGGENTSVGOROLVCLARALLRKTILVYDEA 1456
Db 1381 EEWTSLELAHLKDFVSALPKDLHECAGGENTSVGOROLVCLARALLRKTILVYDEA 1440
Qy 1457 TAAVLDLTDLLIOGTIRTOFEDCTVLIARLNTIMDYTRYVLDKGEIOEYGA PSDLLQ 1516
Db 1441 TAAVLDLTDLLIOGTIRTOFEDCTVLIARLNTIMDYTRYVLDKGEIOEYGA PSDLLQ 1500
Qy 1517 QRGIFYMADAGLV 1531
Db 1501 QRGIFYMADAGLV 1515

RESULT 13
US-10-408-765A-1718
; Sequence 1718, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Baby, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Marnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408.765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1718
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1718

Query Match 72.3%; Score 7037; DB 4; Length 1388;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1387; Conservative 0; Mismatches 1; Indels 42; Gaps 1;

Qy 67 MTPINKTKTALGFLMTIVCWADLFYSPWERSRGIFLAPVFLVSPTLGITTLATFLIQL 126

Db 1 MTPINKTKTALGFLMTIVCWADLFYSPWERSRGIFLAPVFLVSPTLGITTLATFLIQL 60
Qy 127 ERRKVOSSGIMLTFMVALVCAIAIIRSKMTALKEDAOVDLFRDITFYVPSLLIQL 186
Db 61 ERRKVOSSGIMLTFMVALVCAIAIIRSKMTALKE----- 97
Qy 187 VLSCFDSRSLPFSSTIHDPNCPRESSASFLSRIFMWTGLIVRGYQPLEGSDLMSLNK 246
Db 98 -----NCPRESSASFLSRIFMWTGLIVRGYQPLEGSDLMSLNK 138
Qy 247 EDITSEQVVPVLVKNMKKECACTRKQPKVYVSSKDPAQPKESKYDANEAEYALIVKSPQ 306
Db 139 EDITSEQVVPVLVKNMKKECACTRKQPKVYVSSKDPAQPKESKYDANEAEYALIVKSPQ 198
Qy 307 KENPNSLFLKVLKTFGPVFLMSFFPKAIHDIIMFSGQILKLIKPVNDYAPDMQGYFY 366
Db 199 KENPNSLFLKVLKTFGPVFLMSFFPKAIHDIIMFSGQILKLIKPVNDYAPDMQGYFY 258
Qy 367 TVLLFVTRACIOTVLIHQYFHI CFVSGMRITKAVIGAYRRKALVTNRSKSTYGEIYNL 426
Db 259 TVLLFVTRACIOTVLIHQYFHI CFVSGMRITKAVIGAYRRKALVTNRSKSTYGEIYNL 318
Qy 427 MSVDAQRFMDLATYINNIWSAPLOVILALYLLMTLGPSTVAGVAVMWLPVNAVMMAMK 486
Db 319 MSVDAQRFMDLATYINNIWSAPLOVILALYLLMTLGPSTVAGVAVMWLPVNAVMMAMK 378
Qy 487 TKTYQVAMKSKDNRIKLMNEILNGIKVLKYAMELAFKQVLAIRQBELVLRKSAVLS 546
Db 379 TKTYQVAMKSKDNRIKLMNEILNGIKVLKYAMELAFKQVLAIRQBELVLRKSAVLS 438
Qy 547 AVGFTHVCTPPLVALCTFAVYVITDENNILDQTAVASLAFNIIARPLNILEPVYISI 606
Db 439 AVGFTHVCTPPLVALCTFAVYVITDENNILDQTAVASLAFNIIARPLNILEPVYISI 498
Qy 607 VQASVSLKRLRIFLSHEELPDSIERRPVQDGGTNSITVRNATFTARSDPPLNGTF 666
Db 499 VQASVSLKRLRIFLSHEELPDSIERRPVQDGGTNSITVRNATFTARSDPPLNGTF 558
Qy 667 SIPEGALVAVVGQVCGKSSLSALLAEMDKVEGHVAKGSVAVVPOAMVIONDSLENI 726
Db 559 SIPEGALVAVVGQVCGKSSLSALLAEMDKVEGHVAKGSVAVVPOAMVIONDSLENI 618
Qy 727 LFGQLEBPYRYSTIQACALLPDLIELIPSGDRETEIGKGVNLSCGQKORVSLAAVYSNA 786
Db 619 LFGQLEBPYRYSTIQACALLPDLIELIPSGDRETEIGKGVNLSCGQKORVSLAAVYSNA 678
Qy 787 DIYLFDDPLSAVDPAHVGHIFENVIGPKGMLKXKTRILVTHSMGYLPQVDVITYMSGGI 846
Db 679 DIYLFDDPLSAVDPAHVGHIFENVIGPKGMLKXKTRILVTHSMGYLPQVDVITYMSGGI 738
Qy 847 SEMGYSQELLARDGAFAPLRTYASTBOEDAEENGVTGVSQPGKEAKOMENGMLVYDSA 906
Db 739 SEMGYSQELLARDGAFAPLRTYASTBOEDAEENGVTGVSQPGKEAKOMENGMLVYDSA 798
Qy 907 GKOLQRLSSSSSYSGDISRHHNSTAELOKAEKKEETWKLMEADKQGVKLSVMDY 966
Db 799 GKOLQRLSSSSSYSGDISRHHNSTAELOKAEKKEETWKLMEADKQGVKLSVMDY 858
Qy 967 MKAIGLFIPLSLIFLFCNHYASALASNYWLSLMTDDPIVNGQOHTKXRLSVYGALGISQ 1026
Db 859 MKAIGLFIPLSLIFLFCNHYASALASNYWLSLMTDDPIVNGQOHTKXRLSVYGALGISQ 918
Qy 1027 GIAVFGYMAVSIIGILASRCLHVDLHLSILRSPMSFPERTPSGNLVNRFSKELDTVDSM 1086
Db 919 GIAVFGYMAVSIIGILASRCLHVDLHLSILRSPMSFPERTPSGNLVNRFSKELDTVDSM 978
Qy 1087 IPEVIKMFMSLFRNVGACIYILLATPIAIIIPPLGLIYFFVQRFYVASSRQKRLSV 1146
Db 979 IPEVIKMFMSLFRNVGACIYILLATPIAIIIPPLGLIYFFVQRFYVASSRQKRLSV 1038
Qy 1147 SRSPPVSHFNETLIGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLAVRLECV 1206

Db 1039 SRSEVYSHNETLLIGSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANFMLVRLCEV 1098
Qy 1207 GNCIVLFAALFAVISRHSLSAGLVGLSVYSLSQVTTYLNLVMSSEMETNIIVAVERLKE 1266
Db 1099 GNCIVLFAALFAVISRHSLSAGLVGLSVYSLSQVTTYLNLVMSSEMETNIIVAVERLKE 1158
Qy 1267 YSETEKAPWQIOETAPPSWPQVGRVVEFNRYCLRYREDDLPVLRHINVTINGEKVGIV 1326
Db 1159 YSETEKAPWQIOETAPPSWPQVGRVVEFNRYCLRYREDDLPVLRHINVTINGEKVGIV 1218
Qy 1327 GRTAGAGSSLTLLGLFRINESAGEIITIDGINIAKIGLHDLAKFTIIPODPVLSGSLRM 1386
Db 1219 GRTAGAGSSLTLLGLFRINESAGEIITIDGINIAKIGLHDLAKFTIIPODPVLSGSLRM 1278
Qy 1387 NLDPFSQYSDBEVWTSLELAHKDFVSALPKLDHECAGGENTLVSQROLVCLARALLR 1446
Db 1279 NLDPFSQYSDBEVWTSLELAHKDFVSALPKLDHECAGGENTLVSQROLVCLARALLR 1338
Qy 1447 KTKILVLDATVAVDLETTDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTR 1496
Db 1339 KTKILVLDATVAVDLETTDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTR 1388

RESULT 14
US-09-939-853A-88
; Sequence 88, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OR INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939, 853A
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 88
; LENGTH: 1528
; TYPE: PR
; ORGANISM: Mus musculus
US-09-939-853A-88

Query Match 71.9%; Score 7002.5; DB 3; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

Qy 1 MALRFGSADGSDPLMDMNVNTWNTSNPDFTKCFQNTVLVWVPCFYLMACFPPEYFLYLSRH 60
Db 1 MALRFGSADGSDPLMDMNVNTWNTSNPDFTKCFQNTVLVWVPCFYLMACFPPEYFLYLSRH 60
Qy 61 DRGYQIMPLUNKTALGFELMIYVCMADLFYSFWRSGGVLAAPVLVSPITLLGTTLLA 120
Db 61 DRGYQIMPLUNKTALGFELMIYVCMADLFYSFWRSGGVLAAPVLVSPITLLGTTLLA 120
Qy 121 TFLIOLERRKGVQSSGIMLTFWLVALVCALATILRSKIMTALKEDAQVLDLFRDITFYVFS 180
Db 121 TFLIOLERRKGVQSSGIMLTFWLVALVCALATILRSKIMTALKEDAQVLDLFRDITFYVFS 180
Qy 181 LLLIQLVLVSCFSDRSPLFSETIHDNPPCESSASFLSRITFWITGLIVRGYROPLEBSD 240
Db 181 LLLIQLVLVSCFSDRSPLFSETIHDNPPCESSASFLSRITFWITGLIVRGYROPLEBSD 240
Qy 241 LMSINKETTSQVAVVYVKNKKKCAKTRKQPVKVVYS-SKDPAPKSSSKYDANEVEYA 299
Db 241 LMSINKETTSQVAVVYVKNKKKCAKTRKQPVKVVYS-SKDPAPKSSSKYDANEVEYA 299

Qy 300 LTVSPQKEMNPSEFKVLVYTFQGYFLMSFFFKALHDLMPSSGQIILKLKIPVNDYAP 359
Db 301 LTVSPQKEMNPSEFKVLVYTFQGYFLMSFFFKALHDLMPSSGQIILKLKIPVNDYAP 360
Qy 360 DMQGFYTVLLFVYACIOTLVLRHCPFSGRIRKTAIVGAYYRKALVITNSARKST 419
Db 361 DMQGFYTVLLFVYACIOTLVLRHCPFSGRIRKTAIVGAYYRKALVITNSARKST 420
Qy 420 VGEIVNLSVDAQRFMDLTYINNIWGAAPLOVILALYLLMLNLPVSLAGVAVVLMVPV 479
Db 421 VGEIVNLSVDAQRFMDLTYINNIWGAAPLOVILALYLLMLNLPVSLAGVAVVLMVPV 480
Qy 480 NAWNAKTKTYQVAMKSKONRILKMBEILANGIVLKYAMELAFKDKVLAIRQELKVL 539
Db 481 NAWNAKTKTYQVAMKSKONRILKMBEILANGIVLKYAMELAFKDKVLAIRQELKVL 540
Qy 540 KKSAYLAVGTFVWCPELVALCTPAVYVYIDENNILDAQTAFAVSLAFNILEPPLNLT 599
Db 541 KKSAYLAVGTFVWCPELVALCTPAVYVYIDENNILDAQTAFAVSLAFNILEPPLNLT 600
Qy 600 PMVTSIVQASVSLKRLRIFLSHELEBDSIERRPVYDGGGTSITVNAATFTWASDDP 659
Db 601 PMVTSIVQASVSLKRLRIFLSHELEBDSIERRSIKSGEG-NSITVGNATFTWARGBP 659
Qy 660 TLNGITTSIPRGALVAVYVGYGCGKSSLSALLAMBQKVEGHVAIKSVAYVPOAMIQN 719
Db 660 TLNGITTSIPRGALVAVYVGYGCGKSSLSALLAMBQKVEGHVILKGSVAAYVPOAMIQN 719
Qy 720 DLSRENILFGQULEEPPYRSVIOACALLPDEIILPSGDRTEIGEGVNLSSGQKORVSLA 779
Db 720 DLSRENILFGHPLDENYKAWMEACALLPDEIILPSGDRTEIGEGVNLSSGQKORVSLA 779
Qy 780 RAYVSNADIVLFDPLSAVDAHVSKHIFENVIGPKMKLNKTRILVTHSMYSYLPQVDYI 839
Db 780 RAYVSNADIVLFDPLSAVDAHVSKHIFENVIGPKMKLNKTRILVTHGISYLPQVDYI 839
Qy 840 VMSGKISEMWSYOELLARDGAFAELRTYASTQEBDAEENGVTGUSGPGKAKOMENG 899
Db 840 VMSGKISEMWSYOELLARDGAFAELRTYASTQEBDAEENGVTGUSGPGKAKOMENG 896
Qy 900 MLNVDASGKQIOLRSSSSVSGDISRHHNSTAELOKAEKKEFTWKLMEADKQOTGVK 959
Db 900 MLNVDASGKQIOLRSSSSVSGDISRHHNSTAELOKAEKKEFTWKLMEADKQOTGVK 955
Qy 956 LSVYVNDYKALGLTISFLSIFLPMCNHVSALASNYWLSLWTD-DPIVNGTOEHTKVLRSV 1018
Db 956 LSVYVNDYKALGLTISFLSIFLPMCNHVSALASNYWLSLWTD-DPIVNGTOEHTKVLRSV 1015
Qy 1019 YGALGISQGIAVFGYSMAVSTIGLILASRCVLDLHSLTILSPMSGFERTPSGNTLVNRFSK 1078
Db 1016 YGALGISQGIAVFGYSMAVSTIGLILASRCVLDLHSLTILSPMSGFERTPSGNTLVNRFSK 1075
Qy 1079 ELDPVDSMIPRIVIKPMQSLFNVIGACVILLAPPIAIIIPGLIYFPQRPVVASR 1138
Db 1076 ELDPVDSMIPRIVIKPMQSLFNVIGACVILLAPPIAIIIPGLIYFPQRPVVASR 1135
Qy 1139 QKRLLESRSRPVYSHNETLLIGSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANRM 1198
Db 1136 QKRLLESRSRPVYSHNETLLIGSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANRM 1195
Qy 1199 LAVRLCEVGNICIVLFAALFAVISRHSLSAGLVGLSVYSLSQVTTYLNLVMSSEMETNI 1258
Db 1196 LAVRLCEVGNICIVLFAALFAVISRHSLSAGLVGLSVYSLSQVTTYLNLVMSSEMETNI 1255
Qy 1259 VAVERLKEVSETEKAPWQIOETAPPSWPQVGRVVEFNRYCLRYREDDLPVLRHINVTIN 1318
Db 1256 VAVERLKEVSETEKAPWQIOETAPPSWPQVGRVVEFNRYCLRYREDDLPVLRHINVTIN 1315
Qy 1339 GGEKVGIVGRTGAGKSSLTLLGLFRINESAGEIITIDGINIAKIGLHDLAKFTIIPDPV 1378
Db 1316 GGEKVGIVGRTGAGKSSLTLLGLFRINESAGEIITIDGINIAKIGLHDLAKFTIIPDPV 1375
Qy 1379 LFSGSLRMNLDPFSQYSDBEVWTSLELAHKDFVSALPKLDHECAGGENTLVSQROLV 1438

Db 1376 LPSGLRNLDPSQSDSEEWMMALHAKGFVSLPDKLNHECAGEMLNSVGQRLV 1435
Qy 1439 CLAAALRKTKILVLDENTAVIDETDLDLISITRTQEDCTVLTARLNTIMDYTVI 1498
Db 1436 CLAAALRKTKILVLDENTAVIDETDLDLISITRTQEDCTVLTARLNTIMDYTVI 1495
Qy 1499 VLKGEIOEGAPSDLLQORGLFYSMAKAGLV 1531
Db 1496 VLKGEVRECCAPSELLQORGLFYSMAKAGLV 1528

RESULT 15
US-10-618-281-63
Sequence 63, Application US/10618281
Publication No. US20040219609A1
GENERAL INFORMATION:
APPLICANT: Day, Anthony G.
APPLICANT: Esteil, David A.
APPLICANT: Lyons, Eric H.
APPLICANT: Yao, Jian
TITLE OF INVENTION: Methods for Modulating Proteins Not
TITLE OF INVENTION: Previously Known as Proteases
FILE REFERENCE: GC773-2
CURRENT APPLICATION NUMBER: US/10/618,281
PRIORITY FILING DATE: 2003-07-11
PRIORITY FILING DATE: 2002-07-12
NUMBER OF SEQ ID NOS: 92
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 63
LENGTH: 1303
TYPE: PRT
ORGANISM: Homo sapiens
US-10-618-281-63

Query Match 62.3%; Score 6068.5; DB 5; Length 1303;
Best Local Similarity 81.4%; Pred. No. 0;
Matches 1229; Conservative 17; Mismatches 57; Indels 207; Gaps 10;

Qy 22 WNTSNPDTKCFQVTVLWVPCFYLMACFPFYFLYLSHRDGYIOMTPIKTKTALGFL 81
Db 1 WNTSNPDTKCFQVTVLWVPCFYLMACFPFYFLYLSHRDGYIOMTPIKTKTALGFL 60
Qy 82 WIVCWADLFYSFWERSRGIFLAPVPLVSPITLITLTIATLIDERRKGVQSSGIMLTF 141
Db 61 WIVCWADLFYSFWERSRGIFLAPVPLVSPITLITLTIATLIDERRKGVQSSGIMLTF 120
Qy 142 WLVALVCAALATLRKIMTALKEAQAQVDFRDTFYVVSLLIQVLVSCFSDRSPLESET 201
Db 121 WLVALVCAALATLRKIMTALKEAQAQVDFRDTFYVVSLLIQVLVSCFSDRSPLESET 177
Qy 202 IHDNPPRESSASLSRITTFWMTGLYRGYROPLEGSDLSLNKEDTSEGVNPLVKNW 261
Db 178 IHDNPPRESSASLSRITTFWMTGLYRGYROPLEGSDLSLNKEDTSEGVNPLVKNW 237
Qy 262 KKECAKTRKQPVKVVSSKDAQPKSSKVDANEVEALLVKSPOKEMNPSLFVLYKTF 321
Db 238 KKECAKTRN-----SSGSGSCSANTALF-----PA--PTCHKSF 271
Qy 322 GPFLMSPFFKAHIDLMFSGPOLKILKIFVNDTKAPDMQGYFYTVLLFTACGLQTLVL 381
Db 272 QALSLL-----LCRLIKFVNDTKAPDMQGYFYTVLLFTACGLQTLVL 314
Qy 382 HOYHICFVSGMRITKAVTGAIVYRKALVITNSAKKSTVGEIVNIMSVDAORFMDLATYI 441
Db 315 HOYHICFVSGMRITKAVTGAIVYRKALVITNSAKKSTVGEIVNIMSVDAORFMDLATYI 374
Qy 442 NMITSAPLOVTLALYLLMLNGESVLAGVAVMLVMPVNAVAMKTKTYOYAHKSKDNR 501
Db 375 NMITSAPLOVTLALYLLMLNGESVLAGVAVMLVMPVNAVAMKTKTYOYAHKSKDNR 416
Qy 502 IKLNNELLNGIKVLKVLAMBLAFKDKVLAIRQBELKVLKKSAYLSAVGTFTWCTPFLVA 561

Db 417 IKLNNELLNGIKVLKVLAMBLAFKDKVLAIRQBELKVLKKSAYLSAVGTFTWCTPFLVA 476
Qy 562 LCTPAVYVTTDENNIDLAQAFVSLAPNIRPPLNLTIPWYSSIVQASVSKRLRFLS 621
Db 477 LCTPAVYVTTDENNIDLAQAFVSLAPNIRPPLNLTIPWYSSIVQASVSKRLRFLS 528
Qy 622 HEELPSPISERRPVKDDGGTNSIVRNATFTWASDPTLNGITFSPISPEGLVAVVGQV 681
Db 529 GATSERGPWGSRPKHG-----TRQASFSVAPGVLCRPSITFSIPEGALVAVVGQV 581
Qy 682 CGKSLSLALAEWDKVEGHVAIKGSVAYVYPOQAMIONDSIRENLTFCQLEBPYRSVI 741
Db 582 CGKSLSLALAEWDKVEGHVAIKGSVAYVYPOQAMIONDSIRENLTFCQLEBPYRSVI 641
Qy 742 QACALLPDLLETLPSSGDRTEIGEKVNLSSGQKQVSLARAVYSNADITYLPDPLSAVDAA 801
Db 642 QACALLPDLLETLPSSGDRTEIGEKVNLSSGQKQVSLARAVYSNADITYLPDPLSAVDAA 701
Qy 802 VGKIFERNVIGPKMLNKTRILVTHSMSTLPQVDVLIWSSGKISEMGSYQELLARDGA 861
Db 702 VGKIFERNVIGPKMLNKTRILVTHSMSTLPQVDVLIWSSGKISEMGSYQELLARDGA 724
Qy 862 FAEFLRTYASTEQDABENGVTGVSQPKAKQKQNGMLVTDAGKQLOQLSSSSSSYS 921
Db 725 -----S 725
Qy 922 GDISRHHNSTAELQKAEKKEETWKLMEADKAQOQVYLSVYWDYKALGFLSFLSL 981
Db 726 CDL-----QYKLSVYWDYKALGFLSFLSL 753
Qy 982 FMCNHSALASNYLSTMTDPIYNGTQERTKRLSYGALGISQGIYAVRGYSNAVSIGG 1041
Db 754 FMCNHSALASNYLSTMTDPIYNGTQERTKRLSYGALGISQGIYAVRGYSNAVSIGG 813
Qy 1042 ILSRCLHVDLHSLILSPSPFERTPSGNLVNRSFELDPVDSMTEPEVIMKFMGSLFNV 1101
Db 814 ILSRCLHVDLHSLILSPSPFERTPSGNLVNRSFELDPVDSMTEPEVIMKFMGSLFNV 873
Qy 1102 IGACIVILLATPIAIIIPPLGLYFPVORFYVASSRQLRLSEVSRSPPVSHNETLLG 1161
Db 874 IGACIVILLATPIAIIIPPLGLYFPVORFYVASSRQLRLSEVSRSPPVSHNETLLG 933
Qy 1162 VSVTRAEBOERPHOSDLDKVDENQKAYYSIVANRWLAIRLCEVNCIVLPALPAVIS 1221
Db 934 VSVTRAEBOERPHOSDLDKVDENQKAYYSIVANRWLAIRLCEVNCIVLPALPAVIS 993
Qy 1222 RHLISAGLVGSLVSYLOVTVLWLVNRMSSMETNIVAVBRLEKESBTEREAPWJOET 1281
Db 994 RHLISAGLVGSLVSYLOVTVLWLVNRMSSMETNIVAVBRLEKESBTEREAPWJOET 1053
Qy 1282 APPSPQVQVREBRNCLARYEDLPVLHINVTINGEKVGIYGRGTGAKSSLTJGLF 1341
Db 1054 APPSPQVQVREBRNCLARYEDLPVLHINVTINGEKVGIYGRGTGAKSSLTJGLF 1113
Qy 1342 RINESABGEIITDGINIAKIGLHDLRFKTIIPDPVLPFGSLRMLNDPFSQYSDEBEVMT 1401
Db 1114 RINESABGEIITDGINIAKIGLHDLRFKTIIPDPVLPFGSLRMLNDPFSQYSDEBEVMT 1173
Qy 1402 SLELAHLKDFVSALPDLDHECAEGEENLSVGQROVLCARALRKTKIIVLDATAYVD 1461
Db 1174 SLELAHLKDFVSALPDLDHECAEGEENLSVGQROVLCARALRKTKIIVLDATAYVD 1233
Qy 1462 LETDDLIQSTIRTOFEDCTVLTARLNTIMDYTVIVLDKGEIOEGAPSDLLQORGLF 1521
Db 1234 LETDDLIQSTIRTOFEDCTVLTARLNTIMDYTVIVLDKGEIOEGAPSDLLQORGLF 1293
Qy 1522 YSMAKXAGLV 1531
Db 1294 YSMAKXAGLV 1303

Job time : 165.846 secs

```

RESULT 4
US-11-090-439-26
; Sequence 26, Application US/11090439
; Publication No. US20050266442A1
; GENERAL INFORMATION:
; APPLICANT: Squillace, Rachel
; APPLICANT: Weiner, Michael P.
; TITLE OF INVENTION: Immortalized Human Tuberos Sclerosis Null
; TITLE OF INVENTION: Angliomylipoma Cell and Method of Use Thereof
; FILE REFERENCE: 24318-502
; CURRENT APPLICATION NUMBER: US/11/090,439
; CURRENT FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: 60/556,344
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 1581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-090-439-26

```

Query Match 21.6%; Score 2100; DB 7; Length 1581;
Best Local Similarity 31.7%; Pred. No. 7.6e-146;
Matches 524; Conservative 306; Mismatches 567; Indels 254; Gaps 37


```
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 969
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-969

Query Match          10.0%; Score 969; DB 6; Length 410;
Best Local Similarity 50.8%; Pred. No. 9,8e-64;
Matches 197; Conservative 66; Mismatches 121; Indels 4; Gaps 2;

QY 495 MSKDNRIKLMNEILNGIKVLYAMELAFKDKVLAIRQEBLVKKSAVLSAVGTFTWV 554
DB 1 MKQDSRRLTSSILRNKSKITFHGMEGAFIDRVLGITGQELGALRTGGLFVSLSVQ 60
QY 555 CTPELVALCTPAVYVTTIDENNILDAQTA FVSLALFNILRFPNLILPMVTSISIVQASVSLK 614
DB 61 VSTFLVALVFAVHTLVAE-NAMNAEKA FVTLTVLNIINKAQAFIPFSIHSLVQARVSPD 119
QY 615 RLRLFLSHBELEPDSIERPKVYDGGGTSITVRNATFTWARSBDPTLNGITFSIEGALV 674
DB 120 RLVLFLCLEEYDPGVADSSSGSAGKDCITIQSTFFAMSOQSPCLHRINLTVQGGCL 179
QY 675 AVVGQVGKSSLSALAEADKVEGHAIVKGSVAVVPOQAWIQNDSLRNLTFCOLEE 734
DB 180 AVVGQVGKSSLSALAEGLSKVEGVSISGAVALVPOEAVQVTSVENVCFQOEIDP 239
QY 735 PYRSVIOACALLPDLLEILPSSGDRTEIGKEGVNLGGQKQORVSLARAVYSNADIVLPDP 794
DB 240 PMLERVLEACALQPVVDSFPBSIHHSIGQGWNLSGGQKQRLSLARAVYRKAAVLLDDP 299
QY 795 LSAVDAAVGHKIFENVIGPKGMKNKTRILVTHSMYSYLPQVDVVIIVMGSGKISEKSYOE 854
DB 300 LAALDAHVGHQHFNVGIVGPGGILQCTRTLLVTHALHILPOADWIIIVLANGALAEKSGSYOE 359
QY 855 LLARDGAFAEFRTASTTEQEDAEENG 882
DB 360 LLQKQKALVCLL--DQARQPEDRGSGG 384

RESULT 7
US-10-613-744-14
; Sequence 14, Application US/10613744
; Publication No. US20050272093A1
; GENERAL INFORMATION:
; APPLICANT: MacKinnon, Roderick
; APPLICANT: The Rockefeller University
; TITLE OF INVENTION: Assays for Screening Compounds Which Interact With
; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
; TITLE OF INVENTION: Channel Proteins, and Uses Thereof
; FILE REFERENCE: 018512-002901US
; CURRENT APPLICATION NUMBER: US/10/613,744
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US/09/275,252
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: US 09/045,529
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/054,347
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: WO PCT/US99/06307
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 14
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-613-744-14

Query Match          9.1%; Score 883.5; DB 6; Length 407;
Best Local Similarity 48.7%; Pred. No. 1,8e-57;
Matches 167; Conservative 74; Mismatches 95; Indels 7; Gaps 4;

QY 1551 LTRLAEPAPBRY-----RTRRRARFVSKKQNCVAHKNIREQQPLQDVFTLVDKWP 1606
DB 16 LPKQARDL-LPRHSRDRTKRIQIVRKQDCKGNHGNVETRYLTDTITLVLDKMR 74
QY 1607 HTLLIFTPMSPLCSWMLPAMVWMLIAFAHGLDPAPEGTN-VCVTSIHSFSSAPLSIEVQ 1665
DB 75 FNLIIFVWVTVTWLPFGMIMWLLAYIRGDMDDHEDSPWPCVTNNGVSAFLFSIETE 134
QY 1666 VTIGGGRMTEBCEPLAILIIVQNIYGLMNAIMLCIFPKTKQAHRRAETLIFSKAV 1725
DB 135 TTIGGYRVITDKCPREGIILLIQLSVLSGSIYNAFMVGCMPFKISQPKKRARETLVFSTHAV 194
QY 1726 ITLRHGRLCFMLRVGDLRKSMIISATTIMQVVRKTSPEGEVPLHGVDIEMENGVGNG 1785
DB 195 ISMRDGLKCLMFRVGDLSNHTIVASIRAKLIKQKQTSBGEFFILNQTDINVGYYTGDR 254
QY 1786 IFVLAPLIIVHVIDSNSPLVDLAPSDLHHODEIIVLBSGVETGTTTQARTSYLAD 1845
DB 255 LFVISPILSHIENQSPFWFISKAQ-L-PKELEIIVLBEQVATEWGTQCARSSYITSE 313
QY 1846 ILMGQRPVPIAEDGRYSVDYSKFGNTIKVPTPLCARQODE 1888
DB 314 ILMGVRFPVPLTLBDFEYVDYNSFHFETYSTSLSAKELAE 356

RESULT 8
US-10-995-561-967
; Sequence 967, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 967
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-967

Query Match          8.2%; Score 799; DB 6; Length 268;
Best Local Similarity 56.3%; Pred. No. 1,5e-51;
Matches 151; Conservative 56; Mismatches 61; Indels 0; Gaps 0;

QY 1264 LKESYTEKAPWQIQETAPSSWPQVGRVERFRNYCLRYREDLDPVLAHINVTINGEKV 1323
DB 1 MDVYMTPEKAPWRLPTCAQPPWPGQGQIFRDPGLRYRDELPLAVQGVSEFKIHAGBKV 60
QY 1324 GIVGRTGAGKSLTGLFRINESAGEIITIDGINAKIGLMDLRKTIITPQDPVLPSGS 1383
DB 61 GIVGRTGAGKSSLASGLRLQEAAGGIWDGVPILAHGLHLTKLRISIIIPQDPVLPSGS 120
QY 1384 LRMDLPPSOXSDEEVTSLSLAHKDPFVSLAPDLTDHECAEGENLVGQROVLCLARA 1443
DB 121 LRMDLLOHSDDEALMAALETVQAKALVASTPGLQYKCKADRGSDLVGQKQLCLARA 180
QY 1444 LIRKTIIVLDEATAVADLFTDDLLQSTIRQFEDCTVLTIAHRLNTMDYTRVIVLDKG 1503
```

Db 181 LNRKQTLLIDDEATAVDPGTELOQMALGSMFAQCTVLLIARLRASVMDCAVLANMDK 240
QY 1504 EIQEYGAPSDLLQORGLFYSAKADGLV 1531
Db 241 QVAESSPAQLAQKGLFYRLAQESSGLV 268

RESULT 9
US-10-613-744-13

; Sequence 13, Application US/10613744
; Publication No. US20050272093A1
; GENERAL INFORMATION:
; APPLICANT: MacKinnon, Roderick
; TITLE OR INVENTION: Assays for Screening Compounds Which Interact With
; TITLE OR INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
; FILE OR INVENTION: Channel Proteins, and Uses Thereof
; FILE REFERENCE: 018512-002901US
; CURRENT APPLICATION NUMBER: US/10/613,744
; PRIOR FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US/09/275,252
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: US 09/045,529
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/054,347
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: NO PCT/US99/06307
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentln Ver. 2.1
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-613-744-13

Query Match 8.1%; Score 786.5; DB 6; Length 391;
Best Local Similarity 46.2%; Pred. No. 2.2e-50;

Matches 150; Conservative 75; Mismatches 89; Indels 11; Gaps 6;

QY 1564 RTERRAAFVSKGNCNVAAHKNIREQRFL--QDVFTLLVDLKKPHLLIFTSFLCSWL 1621
Db 35 RSRQ-RARLVSKSGRCNIREFGVDAQSRPIFFVDIWTTVLDKKRYMOTVFITAFGLSMF 93
QY 1622 LFLAMWLLIAPAGDL---APGEGINVPCTSTSHSSSAFLSIEVQVITIGGGRNVT 1677
Db 94 LFLGLMVAVVAHVHDLPEFPYPPDNR--PCVENINGMTSAFLSFLSTQVITIGYGFRTVE 151
QY 1678 ECPALILIVONVGMINAMIMGCFEMKTAQAHRAETLISKHAVITLRRHRLCFML 1737
Db 152 QCATALFLFQSLTGLVINSFPGALILAKISRPKKAKTITTSKNAVISKRGGLCLI 211
QY 1738 RVGDLRKSMTISATIHQVVRKTSPEGEVPLHQVDIPMENGVGNGIPLVAPLIYHV 1797
Db 212 RVANLRKSLIGSHIYGLKLTITTPBGETIILDQNNINPVDAENMLFFISLTIYH 271
QY 1798 IDSNSPIYDLAPSDLHHQDLEIIVILGVEVETGTTQARTSYLADEILMGORFVIA 1857
Db 272 IDHNSPFHMAAETL--SQDFELVFLDGTVESTSATCQVRTSYVPEVLWGRFVIVS 330
QY 1858 E-EDGRSYVYSKRGNTIKVPTPLC 1881
Db 331 KTEGKYVDVFNHFKTVEVETPHC 355

RESULT 10

US-10-793-626-326
; Sequence 326, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US

CURRENT APPLICATION NUMBER: US/10/793,626

CURRENT FILING DATE: 2004-03-04

PRIOR APPLICATION NUMBER: 60/164,258

PRIOR FILING DATE: 1999-11-09

NUMBER OF SEQ ID NOS: 4472

SOFTWARE: Patentln Ver. 2.1

SEQ ID NO 326

LENGTH: 548

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: synthetic

OTHER INFORMATION: amino acid sequence

US-10-793-626-326

Query Match 5.3%; Score 514.5; DB 6; Length 548;
Best Local Similarity 27.1%; Pred. No. 3.5e-30;

Matches 157; Conservative 110; Mismatches 211; Indels 101; Gaps 16;

QY 1002 DPLVNGQETKRLSYGALGISQGIYAVGYSMAVSIIGLISRCILHDLHSLASPM 1061
Db 12 DGVIN--NHSLTNOEKFSHGLVAILP-----IFLIVRPP 47
QY 1062 SPFE---RTPSGNLVNRFSKEL-----DTVDSMIPVYIK-----MEMG 1096
Db 48 EPIRQYLAQWTSNKILYDIRKQLYNLQALSVRFYANNQGYISRYINDVEQTKDITLT 107
QY 1097 SLFNVIGACIVILLAPPI-----AAIIIPPLGL--IYFVQRFYVASSRQLRLE 1144
Db 108 GLNMIMWDCITIIILASISIMFPLDVKLFALFIFFPYILTVYFFPGL-----RKLIRVR 162
QY 1145 SVSRSPYVSHFNTLGVSTIRAPESQERFIHQSDLKVDENQAKAYYSIYANRLAVRLE 1204
Db 163 SQALAEVQGFLLHRVQMSYIKSFALIEDNAKQFD--NHNKQFLQRAFOHTRWNAYSFA 219
QY 1205 CV-----GNCIVLFAALFAVISRHSLSAGLVSVSYQVTTYLMVLVMSSEMETNI 1258
Db 220 AINTVTDLGPILYIGVGSYLAIT--GSTVGTTLAFLGYLEQLGFLRLVSSFTTLIQSF 278
QY 1259 VAVRLKEVSETEKE-----APWQIQETAPPSSWPQVGRVFRNYCLAREDDLPVLR 1311
Db 279 ASMDRVQLMDEDDYIDKINGIAQPIKISK-----QIDLKVYSFYNEKEKVLH 328
QY 1312 HINVTINGEKVGVIGRTGKSSLTGLFRINESABGELIIDGINIAKGLHDLRKXIT 1371
Db 329 DINTLTKGRTVAFVGGSGGKSTLINTLIRFYDVYQGEIILIDHNKDFLTGSLRNOIG 388
QY 1372 IIPQDPVLFSGSLRML---DPSQYSDERVTWSLELAHLKDFVSLPDKLDHECAAGE 1428
Db 389 LVQDDNLFSDTYKENLILGRP--DATDDEVEBAKMANAHDFISLNPNGYDTEVGERGV 446
QY 1429 NLSVGQRLVCLARALLRKTILVLDDEATAVADLETDDLIOSTIRIQFEDCTVLLIARL 1488
Db 447 KLSGGQKQRIARIFLNPNPVILDEBATSALDESALIQEALDVLSKORTLLIVAHRL 506
QY 1489 NITMDYTRVIVLDKSGISOEYGAPSDLLQORGLF---YSM 1524
Db 507 SITTHADRIYVWENGRIIVETGTQOLINKGAYEHLYSI 545

RESULT 11

US-10-467-657-1346
; Sequence 1346, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657

```
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1346
; LENGTH: 622
; TYPE: PRF
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1346

Query Match      5.1%; Score 495.5; DB 6; Length 622;
Best Local Similarity 28.7%; Pred. No. 1e-28;
Matches 154; Conservative 106; Mismatches 213; Indels 63; Gaps 13;

QY 1022 LGISGIAVFGYSMAVSIIGGLASRCHLVHDLHSILSRPMSFPFERTPSGNLVNFSKELD 1081
DB 108 LVVIRGICRFTSTYLTMTWVSVMITISKIKDMPAKMLTSSRYHQTPEGTVMNMNLITE 167
QY 1082 TVDSMIEVIMKFMGSLFNVIAGACIVLL-----ATPIAIIIPPLGLIYFVORFYVASS 1137
DB 168 QSVSNASDIFTVLRDITMTVGLTIVLYLWMQSLIYVLMFPLSLT-----SRY---R 220
QY 1138 RQLKRLSVSRSPYSHFN---ETLLGVSVTRAFBEQERFTHQSDL-----KYDE 1184
DB 221 DRLKHVIDSOKSIGTMNVVIAETHQHRVYKLFNGQAQANRPDAVNRITVRLSKITQ 280
QY 1185 NOKAYYP--SIVARNMLAVRLECVGNCIVLPALPAVISRHSLSAGVLGSYSLOYTT 1242
DB 281 ATAAHSPSELSIAIALAV-----VITIALQSQNGYTTTIGEFMAFIYA-MLQNYA 330
QY 1243 YLNMVMSSEMETNIVAVRLKEYSETEKE-----APWOIOETAPPSWPQVGRVER 1296
DB 331 PIKSLANISIPMQTMFLAAGVCAFLDTPPEODKGTALPQRYE-----GRISFR 379
QY 1297 NYCARYEDDPLVLRHINTVINGEKVIGVGTGAGKSLTLGLFRINESABGEIITIGI 1356
DB 380 NVDEYRSDGKIKALDNFNLDIRQGERVALVGRSGSKSTVNLPRFEPEPAGNICIDGI 439
QY 1357 NIAKIGLHDLRFKTIIPQDPLVLFSGSLRMNLDPRSGS-----DEEWTSLSLAHKDF 1411
DB 440 DIADIKDLCAQALVQSQDFLEPDTLFFENV---KRSRPDAGEAELVSLQAAANLSL 495
QY 1412 VSALPDKLDHBCAEGENL SVGQROLVCLARALLRKTKILVDEATAAVDLETDDLIOST 1471
DB 496 IDASPLGLHQBIGSNGSLSGGQGRVAILKDAIILLDETSALDHESEELVQQA 555
QY 1472 IRTQFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDILQOQRLFYSMADQ 1527
DB 556 LERLMENRTGIIVAHRLTTVESADRIIVMDGKIIBOGTHDQLMFQNG-YVTMLRN 610

RESULT 12
US-11-082-389-148
; Sequence 148, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Krogger, Burkhard
; APPLICANT: Schoder, Hartwig
; APPLICANT: Zelder, Oskar
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-131CPCN
; CURRENT APPLICATION NUMBER: US/11/082,389
; PRIOR FILING DATE: 2005-03-16
; PRIOR APPLICATION NUMBER: US 09/603024
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/143262
```

```
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: US 60/151281
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930487.4
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19930489.0
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931549.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931550.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932134.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 446
; SEQ ID NO 148
; LENGTH: 549
; TYPE: PRF
; ORGANISM: Corynebacterium glutamicum
US-11-082-389-148

Query Match      4.8%; Score 468.5; DB 7; Length 549;
Best Local Similarity 26.9%; Pred. No. 8.3e-27;
Matches 141; Conservative 110; Mismatches 201; Indels 73; Gaps 18;

QY 1031 FGYSMAVSIIGGLILASRCHLVHDLHSILSRPMSFPF-----TPSGNLVNFSKELD 1084
DB 50 FGSKLSMEVGR-----DLRSALFGKVNVFSEEREMGQFGABS--LITNINDVQOQ 98
QY 1085 SMIEVIMKFMGSLFNVIAGACI-----VILATPIAIIIPPLGLIYFVORFY 1133
DB 99 MLVQMTSTLMSAPMLAIGIIMAVRODLGLSWMLNVVSIPLIIV--ALIIYRMVPLF 155
QY 1134 VASSRQLKRLSVSRSPYSHFNETLLGVSVTRAF---BEQERFTHOS-----DLKVD 1183
DB 156 QTMOKRIDRIHQIIR-----EOLTGIRIVIRAFVEBDEVREERFTTASKVDADIGVGTG 207
QY 1184 ENQRAYPSIIVANMLAVRLECVGNCIVLPALPAVISRHSLSAGVLGSYSLOYTTY 1243
DB 208 NLMMLMFPAYV---LIMNLSAVA---VIMGAVQVESGET-QIGTLPAFLQYIMQILMG 259
QY 1244 LNMVMSSEMETNIVAVRLKEYSETEKEAPWOIOET-APPSSWPQVGRVERFYCYRY 1302
DB 260 VMMAAFPMVWVRAVSAADRIGEVLETPSY--QAPETPAQPT--SAGEIVFNNAFTAY 315
QY 1303 REDIDPVLRHINTVINGEKVIGVGTGAGKSLTLGLF-RINESABGEIITIDGINAKI 1361
DB 316 PGADDPVLMNVSPFVADGSTTAIIIGTSGSGKTTTL-IGLVPRLPVTEGDVTVDDVDAEF 374
QY 1362 GLHDLRFKTIIPQDPLVLFSGSLRMNLDPRFQSGSDE-----EWTSLSLAHKDFVSLP 1416
DB 375 EPLKMDRIGLVPKSFLFSGTIASNL---RYGNEDATETQWOLALIAQADAPVREMP 430
QY 1417 DKLDECAEGENL SVGQROLVCLARALLRKTKILVDEATAAVDLETDDLIOGTIRTOF 1476
DB 431 EGLDSEIAQSGTIVSGGQGRVAILKDAIILLDETSALDHESEELVQQA 490
QY 1477 EDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDILQOQRLF 1521
DB 491 PDATKLIYAGVSTIRBADQIVVDNNGEVVIGITHYMLNMGTY 535

RESULT 13
US-10-995-561-575
; Sequence 575, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
```

```
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 575
LENGTH: 701
TYPE: PRT
ORGANISM: Homo sapiens
US-10-995-561-575

Query Match
Best Local Similarity 25.2%; Score 419.5; DB 6; Length 701;
Matches 149; Conservative 130; Mismatches 234; Indels 79; Gaps 21;

QY 977 LSTFLFM--CNHVSALASNYWLSLMTDDPIVNGTQHTKRLSVYGALGISQGIAYGYS 1034
DB LSTFLVLVLTSSLGEMALPFTGRITWILQDGSADFTTRNLITMSILITASVLEF--- 196
QY 1035 MAVSIG-GILASRCLHV-----DLHSILRSPMSFEPTPSGNLVNRFKELDVDSMI 1087
DB 197 ----VGDCIYNNMNGHSHLQGEVFAVLROETFEFOQNOTGNIMSKVETDSTLSDSL 252
QY 1088 PEVIMKMGSLFENVIGACIV-ILT-----ATPIAIIIPPLGIYFFVQRFYVASSROLK 1141
DB 253 SENISLFLMWL--VRGCLLIGIMMGSVSLMTVTLITPLFLPKKVKKYYQLLEVOVR 310
QY 1142 RLSEVSRSPTYSHNETLLGVSVIRAFEEQERFIHQSDLVKDE---NOK--AYPSIV 1194
DB 311 --ESLASSOVA--TEALSMPPTVRSFANEGEAQKFEKLOEIKTLNQEAVAY---A 362
QY 1195 ANRW-----LAVRLCEVNCIVLPAALFAVISRHSLSAGLVGSYSLOVTTYIANW 1246
DB 363 VNSMTTISIGMLKLVGILYIGQOLVTSGA-----VSSGNLVTVLYQMOFTQAVEV 413
QY 1247 LVNRSSEMETNIVAVERLKEYSETEKEAPMOIOETAPSS---WPQGRVFEFNNYCLRY 1302
DB 414 LLSYPRVQKAVSGSEKFEYLDRTPR-----CPSSGLLPPLHLEGLVQFQDVSPAY 465
QY 1303 --REDDLPLRHINVTNGEKGVIQVGTGAKSSLTGLGRINESAGEIIIDGINIAK 1360
DB 466 PNRBDV-LVLOGLFTLTPRGEVTLVGPNSGKSTVALLQNLVQPTGGQLLDGKPLPQ 524
QY 1361 IGLHDLRFKTIIPQDPVLFSGSLRML--DPFSQYSEDEWTSLELAHKDFPSALPDK 1418
DB 525 YEHRYLHRQVAAGQEPQVGRSLOENIAYGLTQKPTMEETITAAVNSGAHSFISGLPQG 584
QY 1419 LDHECAGGENLSVGOROLVCARALLRKTKILVLDATAVD---LETDDLIOSTIRT 1474
DB 585 YDTEVDEAGSOLSGGQQAVALARALKRCPVLILDDATSLDANSQLOVEQLLYES--P 642
QY 1475 QFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDILQORGLFYSPMAK 1526
DB 643 ERYRSRVLLITQHLSTVEQADHILFLEGAIREGGTHQOLMEKKGCYMAWQ 694

RESULT 14
US-10-995-561-574
Sequence 574, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
FILE REFERENCE: CL001559
DETECTION AND USES THEREOF
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 574
LENGTH: 808
TYPE: PRT
```

```
ORGANISM: Homo sapiens
US-10-995-561-574

Query Match
Best Local Similarity 4.3%; Score 419.5; DB 6; Length 808;
Matches 149; Conservative 130; Mismatches 234; Indels 79; Gaps 21;

QY 977 LSTFLFM--CNHVSALASNYWLSLMTDDPIVNGTQHTKRLSVYGALGISQGIAYGYS 1034
DB 247 LSTFLVLVLTSSLGEMALPFTGRITWILQDGSADFTTRNLITMSILITASVLEF--- 303
QY 1035 MAVSIG-GILASRCLHV-----DLHSILRSPMSFEPTPSGNLVNRFKELDVDSMI 1087
DB 304 ----VGDCIYNNMNGHSHLQGEVFAVLROETFEFOQNOTGNIMSKVETDSTLSDSL 359
QY 1088 PEVIMKMGSLFENVIGACIV-ILT-----ATPIAIIIPPLGIYFFVQRFYVASSROLK 1141
DB 360 SENISLFLMWL--VRGCLLIGIMMGSVSLMTVTLITPLFLPKKVKKYYQLLEVOVR 417
QY 1142 RLSEVSRSPTYSHNETLLGVSVIRAFEEQERFIHQSDLVKDE---NOK--AYPSIV 1194
DB 418 --ESLASSOVA--TEALSMPPTVRSFANEGEAQKFEKLOEIKTLNQEAVAY---A 469
QY 1195 ANRW-----LAVRLCEVNCIVLPAALFAVISRHSLSAGLVGSYSLOVTTYIANW 1246
DB 470 VNSMTTISIGMLKLVGILYIGQOLVTSGA-----VSSGNLVTVLYQMOFTQAVEV 520
QY 1247 LVNRSSEMETNIVAVERLKEYSETEKEAPMOIOETAPSS---WPQGRVFEFNNYCLRY 1302
DB 521 LLSYPRVQKAVSGSEKFEYLDRTPR-----CPSSGLLPPLHLEGLVQFQDVSPAY 572
QY 1303 --REDDLPLRHINVTNGEKGVIQVGTGAKSSLTGLGRINESAGEIIIDGINIAK 1360
DB 573 PNRBDV-LVLOGLFTLTPRGEVTLVGPNSGKSTVALLQNLVQPTGGQLLDGKPLPQ 631
QY 1361 IGLHDLRFKTIIPQDPVLFSGSLRML--DPFSQYSEDEWTSLELAHKDFPSALPDK 1418
DB 632 YEHRYLHRQVAAGQEPQVGRSLOENIAYGLTQKPTMEETITAAVNSGAHSFISGLPQG 691
QY 1419 LDHECAGGENLSVGOROLVCARALLRKTKILVLDATAVD---LETDDLIOSTIRT 1474
DB 692 YDTEVDEAGSOLSGGQQAVALARALKRCPVLILDDATSLDANSQLOVEQLLYES--P 749
QY 1475 QFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDILQORGLFYSPMAK 1526
DB 750 ERYRSRVLLITQHLSTVEQADHILFLEGAIREGGTHQOLMEKKGCYMAWQ 801

RESULT 15
US-11-105-268-53
Sequence 53, Application US/11105268
Publication No. US20050260204A1
GENERAL INFORMATION:
APPLICANT: Allan, Christian
TITLE OF INVENTION: ANTI-IL-9 ANTIBODY FORMULATIONS AND USES THEREOF
FILE REFERENCE: 10271-126-999
CURRENT APPLICATION NUMBER: US/11/105,268
CURRENT FILING DATE: 2005-04-12
PRIORITY FILING DATE: 2004-04-12
NUMBER OF SEQ ID NOS: 60
SOFTWARE: PatentIn version 3.2
SEQ ID NO 53
LENGTH: 808
TYPE: PRT
ORGANISM: Homo sapiens
US-11-105-268-53

Query Match
Best Local Similarity 4.3%; Score 419.5; DB 7; Length 808;
Matches 149; Conservative 130; Mismatches 234; Indels 79; Gaps 21;

QY 977 LSTFLFM--CNHVSALASNYWLSLMTDDPIVNGTQHTKRLSVYGALGISQGIAYGYS 1034
```

Db 247 LSLFLVLVLSLSLGMALPFTGLTDLWLDGSDADFTFRNLTLMSILTLIASAVLEF--- 303
QY 1305 MAVSIG-GILASRCLHV-----DLHSLRSPMSFPERTPSGNLVNRFSKELDTVDSMI 1087
Db 304 ---VGDCIYNNNTMGHSHLOGEVGAFLROETEFPOQNGNTMSKRVTEDTSTLSDSL 359
QY 1088 PEVIKMFWSLFNVIGACIV-ILL-----ATPAAIIPPLGLIYFVQRFYVASSRQJK 1141
Db 360 SENISLFLMWL--VRGLCLGLIMLWGSVLNVTLITLPLFLPKVKGMVQLLEVCVR 417
QY 1142 RLESVSRSPVYSHFETLLGVSIVAPFEQERFIHOSDKYDE---NOK--AYPSIV 1194
Db 418 --ESLAKSSQVA--IEALSAMPYVRSFANEBGEAQFREKLOEIKTLNQKEAVAY---A 469
QY 1195 ANRW-----LAVLECVGNCIVLFAALFAVISRHSISAGLVGLSVSYSLQVTTYLNM 1246
Db 470 VNSWTTISGMLKKGIYIGQLVTSQA-----VSSGNLVTFVLVQMOTQAVEV 520
QY 1247 LVRRSSEMETNI VAVERLKEYSETEKEAPWQIOETAPPS---WPVGRVYEFRRYCLRY 1302
Db 521 LLSIYPRVQKAVGSSEKIFEVYLDRTPR-----CPPSGLLTPLHLEGLVQFDVSPAY 572
QY 1303 --REDLDVLAHINTVINGEYVIGRTGAGKSLTLGLFRINESAGEIITDGINIAK 1360
Db 573 PNRPDV-LVLOGLTFTLRPGEVTLVGPNGSGKSTVAALLQNLXQPTGQQLLDGKPLPQ 631
QY 1361 IGLHDLRFKTIITIPQDPVLFSGSLRML--DPFSQYSDSEVWTSLELAHLKDFVSALPDK 1418
Db 632 YEHRYLHRQVANAAGDEPOVFGKSLQENIAYGLTQKPTMEBITAANA VKSGANHSFISGLPQG 691
QY 1419 LDHBCAEGGENLSVGQRLVCLARALLRKTKILVLDEATAVD---LETDDLIOSTIRT 1474
Db 692 YDTEVDEAGSQLSGGQRLVALRALIRKPCVLILDDATSA LDANSQLOVEQLYES--P 749
QY 1475 QPEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGA PSDLLQQRGLFYSMAX 1526
Db 750 ERYSRSVLLITQHLSLVEQADHILFLREGAIRGGTHQOLMEKGCYMAWVQ 801

Search completed: December 15, 2005, 15:50:09
Job time : 11.1294 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:27:22 ; Search time 173.071 Seconds
(without alignments)
4800.712 Million cell updates/sec

Title: US-10-665-283-4

Perfect score: 9734

Sequence: 1 MAARGFCSAGSDPLMDMNV.....NTIKVPTPLCTARQLDSDRS 1891

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_21: *
1: geneseqp1980s: *
2: geneseqp1990s: *
3: geneseqp2000s: *
4: geneseqp2001s: *
5: geneseqp2002s: *
6: geneseqp2003as: *
7: geneseqp2003bs: *
8: geneseqp2004s: *
9: geneseqp2005s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	9734	100.0	1891	9	ADY86941 Human MRP
2	9734	100.0	1927	9	ADY86943 MRP1-mous
3	9734	100.0	1927	9	ADY86938 Human MRP
4	9704	99.7	1947	9	ADY86945 Human MRP
5	7860	80.7	1531	2	AAW57486 Human MRP
6	7860	80.7	1531	2	AAW74471 Human mul
7	7860	80.7	1531	2	AAW98994 Human mul
8	7860	80.7	1531	3	AAV55799 Human mul
9	7860	80.7	1531	3	AAV78873 Multidrug
10	7860	80.7	1531	5	ABG61810 Prostate
11	7860	80.7	1531	6	ABM35012 Cancer ba
12	7860	80.7	1531	6	ADB20865 MRP1 base
13	7860	80.7	1531	7	ADB87954 Human UGT
14	7860	80.7	1531	7	ADB96937 Human MDR
15	7860	80.7	1531	7	ADB92128 Human MDR
16	7860	80.7	1531	7	ADD44764 Human pro
17	7860	80.7	1531	8	ADN97111 MRP1 prot
18	7860	80.7	1531	8	ADU06426 Novel bro
19	7860	80.7	1531	9	ADX06202 Cyclin-de
20	7856	80.7	1530	7	ADN95929 Human BEC
21	7849	80.6	1531	2	AAW54928 Multidrug
22	7849	80.6	1531	2	AAW93153 Multi-dru
23	7849	80.6	1531	2	AAW57485 Human mul
24	7849	80.6	1531	2	AAW74470 Human mul

25	7849	80.6	1531	2	AAW98893 Human mul
26	7849	80.6	1531	3	AAV55798 Human mul
27	7849	80.6	1531	3	AAV78872 Human mul
28	7849	80.6	1531	3	AAB03582 Multidrug
29	7763	79.8	1515	8	ADU24094 Human Cys
30	7610	78.2	1489	2	AAW96952 Multi-dru
31	7587.5	77.9	1482	8	ADN03902 Antipsoi
32	7587.5	77.9	1482	8	ADP23422 PRO poly
33	7572	77.8	1481	5	ABP52108 Homo sapi
34	7560.5	77.7	1482	2	AAW96953 Multi-dru
35	7535	77.4	1475	9	ADX06206 Cyclin-de
36	7510.5	77.2	1472	9	ADX06204 Cyclin-de
37	7500.5	77.1	1472	2	AAW96954 Multi-dru
38	7230.5	74.3	1530	9	ADP56611 Bovine MR
39	7195.5	73.9	1416	9	ADX06208 Cyclin-de
40	7194	73.9	1417	2	AAW96955 Multi-dru
41	7037	72.3	1388	7	ADJ69912 Human hea
42	7002.5	71.9	1528	2	AAW57487 Murine mu
43	7002.5	71.9	1528	2	AAW74472 Mouse mul
44	7002.5	71.9	1528	2	AAW9895 Mouse mul
45	7002.5	71.9	1528	3	AAV55800 Murine mu

ALIGNMENTS

RESULT 1					
ID	ADY86941	standard; protein; 1891 AA.			
XX					
AC	ADY86941;				
DT	02-JUN-2005	(first entry)			
XX					
DE	Human MRP1-Kir6.2 mutant delc36 fusion protein, SEQ ID NO: 4.				
XX					
KW	Ionophore; biosensor; drug screening; diagnostic;				
KW	microorganism detection; potassium channel; fusion protein;				
XX	multidrug resistance protein 1; MRP1; Kir6.2; mulein.				
OS	Homo sapiens.				
OS	Mus musculus.				
XX					
CH	Chimeric.				
XX					
FT	Key	Location/Qualifiers			
FT	Region	1..1531			
FT	/note= "Multidrug resistance protein 1 (MRP1)"	1532..1537			
FT	Region	/note= "Hexaglycine spacer"			
FT	Region	1538..1891			
FT	Region	/note= "Kir6.2 mutant delc36 protein"			
XX					
XX	US2005063989-A1.				
XX	24-MAR-2005.				
XX	22-SEP-2003; 2003US-00665283.				
XX	22-SEP-2003; 2003US-00665283.				
XX	(COMS) COMMISSARIAT ENERGIE ATOMIQUE.				
XX	Derand R, Garcia E, Prost A, Revilloud J, Vivaudou M;				
XX	WPI; 2005-252611/26.				
XX	New ion channel hybrid protein, used as electrical sensor for screening				
XX	an agonist/antagonist of a membrane protein and for detecting a				
XX	contaminant/pollutant in a sample.				
XX	Claim 25; SEQ ID NO 4; 78pp; English.				
XX	The present invention relates to a hybrid protein consisting essentially				

CC of the fusion of a membrane protein with an ion channel which is not
CC naturally coupled to the membrane protein. The hybrid protein is used as
CC an electrical sensor for screening of an agonist/antagonist of a membrane
CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC invention is also useful for medical diagnostics and microanalysis. The
CC detection. The present sequence is a fusion protein comprising human
CC multimeric resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel Kir6.2 mutant detac36 protein.
XX
SQ Sequence 1891 AA;
Query Match 100.0%; Score 9734; DB 9; Length 1891;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCADSDSDPLMDNNVTWNTSNPDFTKCFQNTVLVWVPCFYLMACPPFFLYLSRH 60
Db 1 MALRGFCADSDSDPLMDNNVTWNTSNPDFTKCFQNTVLVWVPCFYLMACPPFFLYLSRH 60
QY 61 DRGYIQMTPLNKKTALGFLIMVCMADLPYSFWERSRGIFLAPVFLVSPILLGITTLA 120
Db 61 DRGYIQMTPLNKKTALGFLIMVCMADLPYSFWERSRGIFLAPVFLVSPILLGITTLA 120
QY 121 TELLOERKGVQSSGIMLTFWVALVCAALIRSKITMALKEDQVDLPFDITFFYYS 180
Db 121 TELLOERKGVQSSGIMLTFWVALVCAALIRSKITMALKEDQVDLPFDITFFYYS 180
QY 181 LLLIQLVLSCFSDSPLESETIHDNPPCPRESSASFLSRITFMWITGLIVRGYROPLEBSD 240
Db 181 LLLIQLVLSCFSDSPLESETIHDNPPCPRESSASFLSRITFMWITGLIVRGYROPLEBSD 240
QY 241 LMSLNKEDTSQVVPVLVKNKKCECAKTRKOPVKKVYSSKDPAPCKESSKYDAVEEVAL 300
Db 241 LMSLNKEDTSQVVPVLVKNKKCECAKTRKOPVKKVYSSKDPAPCKESSKYDAVEEVAL 300
QY 301 IVKSPKEMNSLPEFVLVYKTEGPFYLMSEFFRAIHDLMFSGPOLIKLIKFNVDTKAPD 360
Db 301 IVKSPKEMNSLPEFVLVYKTEGPFYLMSEFFRAIHDLMFSGPOLIKLIKFNVDTKAPD 360
QY 361 WQGFYFVTLVFTVACTQTLVHOYFHIQFVSGMRITKTAIVTGAIVYKALVITNSARKSTV 420
Db 361 WQGFYFVTLVFTVACTQTLVHOYFHIQFVSGMRITKTAIVTGAIVYKALVITNSARKSTV 420
QY 421 GEIVNLSVDAQRPMDLATYINMISAPLOVILALYLLMLNGPSVLGAVAVMLVMEVN 480
Db 421 GEIVNLSVDAQRPMDLATYINMISAPLOVILALYLLMLNGPSVLGAVAVMLVMEVN 480
QY 481 AVMAKTKTYOVAMHKSQDNRIKLMNELNGIKVLKYAMELAFQDKVLAROEELKYLK 540
Db 481 AVMAKTKTYOVAMHKSQDNRIKLMNELNGIKVLKYAMELAFQDKVLAROEELKYLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENMILDAQTAFFVSLALFNILRPFLNLP 600
Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENMILDAQTAFFVSLALFNILRPFLNLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPYKGGTNSITVRNATFTWASDPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPYKGGTNSITVRNATFTWASDPT 660
QY 661 LNGTHFTSPREGALVAVVQVCGKSSLSALLAENDKVEGHVAILGSAVYVPOQAMIOND 720
Db 661 LNGTHFTSPREGALVAVVQVCGKSSLSALLAENDKVEGHVAILGSAVYVPOQAMIOND 720
QY 721 SLRENILFGCOLLEPYRSVIOACALPDLEILPSGRTIEGKVNLSGGOKORVSLAR 780
Db 721 SLRENILFGCOLLEPYRSVIOACALPDLEILPSGRTIEGKVNLSGGOKORVSLAR 780
QY 781 AVYSNADIIYLPDPLSAVDAVGHKIFENVIQPKMELNKKTRILVTHSMYSYLPQVDVIV 840
Db 781 AVYSNADIIYLPDPLSAVDAVGHKIFENVIQPKMELNKKTRILVTHSMYSYLPQVDVIV 840
QY 841 MSGKISMGSYOEILARDGAPAEFLRTYASTBOQDAEENGVTGVSQPGKAEAKOMENGM 900
Db 841 MSGKISMGSYOEILARDGAPAEFLRTYASTBOQDAEENGVTGVSQPGKAEAKOMENGM 900

Db 841 MSGKISMGSYOEILARDGAPAEFLRTYASTBOQDAEENGVTGVSQPGKAEAKOMENGM 900
QY 901 LVYDSAGKOLQROUSSSSSYSGDISRHNNSTABLOKAEKEETWKLMEADKAQTQVYL 960
Db 901 LVYDSAGKOLQROUSSSSSYSGDISRHNNSTABLOKAEKEETWKLMEADKAQTQVYL 960
QY 961 SVYWDYKAIQGLFISPLISFLPWCNHYSAIASNYWLSLMTDDPIVNGTOEHTKRLSYVG 1020
Db 961 SVYWDYKAIQGLFISPLISFLPWCNHYSAIASNYWLSLMTDDPIVNGTOEHTKRLSYVG 1020
QY 1021 ALGISOGIAVNGYMAVSIIGIILASRCLHVDLHSLIRSPSPFEPTPSGNLVNRFSEL 1080
Db 1021 ALGISOGIAVNGYMAVSIIGIILASRCLHVDLHSLIRSPSPFEPTPSGNLVNRFSEL 1080
QY 1081 DTVDSMIPEVIMKMGSLFNVIGACIYILLATPIAIIIPPLGIYEFVQGFYVASSROL 1140
Db 1081 DTVDSMIPEVIMKMGSLFNVIGACIYILLATPIAIIIPPLGIYEFVQGFYVASSROL 1140
QY 1141 KRLSVSRSPYSHFNFTLLGVSYTRAFPEOERTIHOSDLKVDENQAKYTPSIYANRWLA 1200
Db 1141 KRLSVSRSPYSHFNFTLLGVSYTRAFPEOERTIHOSDLKVDENQAKYTPSIYANRWLA 1200
QY 1201 VRLCVCNCTVLPAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVRSSEMETNIVA 1260
Db 1201 VRLCVCNCTVLPAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVRSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMOIETAPPSSWPQVGRVFRNYCJRYREDLPVLRHINVTNGG 1320
Db 1261 VERLKEYSETEKEAPMOIETAPPSSWPQVGRVFRNYCJRYREDLPVLRHINVTNGG 1320
QY 1321 EKVGIVGRTGAKSSLTGLFRIINESABGIIIDGINIAKIGLHDLAPKTIITIQDDPLF 1380
Db 1321 EKVGIVGRTGAKSSLTGLFRIINESABGIIIDGINIAKIGLHDLAPKTIITIQDDPLF 1380
QY 1381 SCSLRNMLDPPSOYSDDEWMTSLKLAHKDPVSALPKLDHECAGEGNSLVGQROLVCL 1440
Db 1381 SCSLRNMLDPPSOYSDDEWMTSLKLAHKDPVSALPKLDHECAGEGNSLVGQROLVCL 1440
QY 1441 ARALLRKTILVDEATAVVDLETDLDIOSTIRTOFEDCTVLTIAHRLNTIMDTRYVL 1500
Db 1441 ARALLRKTILVDEATAVVDLETDLDIOSTIRTOFEDCTVLTIAHRLNTIMDTRYVL 1500
QY 1501 DKGEIOEYGAUSDLLQORGLFYMAKODAGLVGGGGGMLSRKGIIPREYVLTSLAEDPAE 1560
Db 1501 DKGEIOEYGAUSDLLQORGLFYMAKODAGLVGGGGGMLSRKGIIPREYVLTSLAEDPAE 1560
QY 1561 PRYTRRRRARFVSKKNCNVAHKNIREOGRFLODVFETTLVDLKWPHTLITFMSPLCSW 1620
Db 1561 PRYTRRRRARFVSKKNCNVAHKNIREOGRFLODVFETTLVDLKWPHTLITFMSPLCSW 1620
QY 1621 LLFAMVWMLIAFAHGDAPGEGTVPCVTSIHSFSAFLFSEIVQVITGFQGRVTEBCP 1680
Db 1621 LLFAMVWMLIAFAHGDAPGEGTVPCVTSIHSFSAFLFSEIVQVITGFQGRVTEBCP 1680
QY 1681 LAIILIVONIVGMINAIVMGCIFMKTQAQHRRAETLIESKHAVITLJBHRLCFMRLRVG 1740
Db 1681 LAIILIVONIVGMINAIVMGCIFMKTQAQHRRAETLIESKHAVITLJBHRLCFMRLRVG 1740
QY 1741 DLKRSMTIASITHQVVRKTTSPGEVAPPLHQVDIPENGGGNGIFLVAALIIYHYIDS 1800
Db 1741 DLKRSMTIASITHQVVRKTTSPGEVAPPLHQVDIPENGGGNGIFLVAALIIYHYIDS 1800
QY 1801 NSPLDYLDAPSDIAHHODLEIIVILEGVEETGTTQARTSYLADBIIMGQRFVPIVAED 1860
Db 1801 NSPLDYLDAPSDIAHHODLEIIVILEGVEETGTTQARTSYLADBIIMGQRFVPIVAED 1860
QY 1861 GRYSVDYSKFGNTIKVPTPLCTARQLEDERS 1891
Db 1861 GRYSVDYSKFGNTIKVPTPLCTARQLEDERS 1891
RESULT 2
ADY86943

ID ADY86943 standard; protein; 1927 AA.
XX
AC ADY86943;
XX
DT 02-JUN-2005 (first entry)
XX
DE MRP1-mouse K1r6.2 mutant (KR370AA) fusion protein, SEQ ID NO: 6.
XX
KW Ionophore; biosensor; drug screening; diagnostic;
KW microorganism detection; potassium channel; fusion protein;
KW multimeric resistance protein 1; MRP1, K1r6.2; murein.
XX
OS Homo sapiens.
OS Mus musculus.
OS Chimeric.
XX
FH Key
FH Region
FT Location/Qualifiers
FT 1..1531
FT /note= "Multidrug resistance protein 1 (MRP1)"
FT 1532..1537
FT /note= "Hexaglycine spacer"
FT 1538..1927
FT /note= "K1r6.2 mutant protein"
FT MISC-difference 1907
FT /note= "Wild-type Lys substituted by Ala"
FT MISC-difference 1908
FT /note= "Wild-type Arg substituted by Ala"
XX
PN US2005063989-A1.
XX
PD 24-MAR-2005.
XX
PF 22-SEP-2003; 2003US-00665283.
XX
PR 22-SEP-2003; 2003US-00665283.
XX
PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
XX
PI Derand R, Garcia E, Prost A, Revilloud J, Vivaudou M,
XX
DR WPI; 2005-252611/26.
XX
XX
PT New ion channel hybrid protein, used as electrical sensor for screening
PT an agonist/antagonist of a membrane protein and for detecting a
PT contaminant/pollutant in a sample.
XX
PS Claim 25; SEQ ID NO 6; 78bp; English.
XX
CC The present invention relates to a hybrid protein consisting essentially
CC of the fusion of a membrane protein with an ion channel which is not
CC naturally coupled to the membrane protein. The hybrid protein is used as
CC an electrical sensor for screening of an agonist/antagonist of a membrane
CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multimeric resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel K1r6.2 mutant (KR370AA) protein.
XX
SQ Sequence 1927 AA;
XX
Query Match 100.0%; Score 9734; DB 9; Length 1927;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 121 TFLIQERRKGVQSSGIMLTFMVALVCALAILRSKIMTALKEDAQVDLPDITFYVYS 180
181 LLLIQLVLSGFSRDRPLFSETHDNPDPSSASFLSRITFMWITGLIVRGYRPLEGSD 240
181 LLLIQLVLSGFSRDRPLFSETHDNPDPSSASFLSRITFMWITGLIVRGYRPLEGSD 240
241 LMSLNKEDTSEQVVPVLRKWKKECATRQPVVYVSSQDPAPKSSSKVDANEVEAL 300
241 LMSLNKEDTSEQVVPVLRKWKKECATRQPVVYVSSQDPAPKSSSKVDANEVEAL 300
301 IYKSPQKEMPSPLEKVLVYKTFGPFLMSPFFKALHDMFSGPOILKLLIKFVNDTPAD 360
301 IYKSPQKEMPSPLEKVLVYKTFGPFLMSPFFKALHDMFSGPOILKLLIKFVNDTPAD 360
361 WQGYFTVLLFVTAQTLVLAHQVFIQVSGMRKTAIVAGVYRKALVITNSARKSTV 420
361 WQGYFTVLLFVTAQTLVLAHQVFIQVSGMRKTAIVAGVYRKALVITNSARKSTV 420
421 GEIVNLSVDAQRPMDLATYINMWSAPLOYIALYILNMLGRSVLAGVAVMPLVYN 480
421 GEIVNLSVDAQRPMDLATYINMWSAPLOYIALYILNMLGRSVLAGVAVMPLVYN 480
481 AVAMAKTKTYQVAHMKSKNRKILMNEILNGIKVLKYAMELAFKDKVLAIROBELKYLK 540
481 AVAMAKTKTYQVAHMKSKNRKILMNEILNGIKVLKYAMELAFKDKVLAIROBELKYLK 540
541 KSAVLSAVGTFTWVCTPFLVALCTFAVYVITIDENNILDAQTAFAVSLAFNLRPEPLILP 600
541 KSAVLSAVGTFTWVCTPFLVALCTFAVYVITIDENNILDAQTAFAVSLAFNLRPEPLILP 600
601 MVISSIVQASVSLKRLIFLSHELEBDSIERRPVKDGGGTSITVANAFTWARSDDPT 660
601 MVISSIVQASVSLKRLIFLSHELEBDSIERRPVKDGGGTSITVANAFTWARSDDPT 660
661 LMGITFSIPRGALVAVVGQVCGKSSLSLALBMDVBEHVAKGSVAVPOQAMQND 720
661 LMGITFSIPRGALVAVVGQVCGKSSLSLALBMDVBEHVAKGSVAVPOQAMQND 720
721 SLRENILFGCQLEPEYRYSVIOACALLPDEILPDSGRTEIGEGVNLSGGOKRVSILAR 780
721 SLRENILFGCQLEPEYRYSVIOACALLPDEILPDSGRTEIGEGVNLSGGOKRVSILAR 780
781 AVYSNADIVLPDPLSANDAVHGVKHIFENYIGPKMLKNTKRLIVTHSMGYLPQVDYIIV 840
781 AVYSNADIVLPDPLSANDAVHGVKHIFENYIGPKMLKNTKRLIVTHSMGYLPQVDYIIV 840
841 MSGGKISEMGSYOELLARDGAFAPLRTVASTQBOBAENGVTGVSQPKKAKOMENGM 900
841 MSGGKISEMGSYOELLARDGAFAPLRTVASTQBOBAENGVTGVSQPKKAKOMENGM 900
901 LVYDSAGKQOLQRLSSSSSYSGDISRHHNSTAEIOKAACKETWKLMEADKAQOTGVKL 960
901 LVYDSAGKQOLQRLSSSSSYSGDISRHHNSTAEIOKAACKETWKLMEADKAQOTGVKL 960
961 SVTWIDYKAIGLFISPLSTFLMGNVSAVSAVNTSLMTDDPIVNGTOHTYRSLSVYG 1020
961 SVTWIDYKAIGLFISPLSTFLMGNVSAVSAVNTSLMTDDPIVNGTOHTYRSLSVYG 1020
1021 ALGISGIAVFGVSMVSTGGILASRCILHVDLHSLRSPMSFEPSTPSGNLVNRPSKEL 1080
1021 ALGISGIAVFGVSMVSTGGILASRCILHVDLHSLRSPMSFEPSTPSGNLVNRPSKEL 1080
1081 DTVDSMIPEVIAKMFMSLFPNVIGACIVILATPAAIIIPPLGLIYFVQRFVASSRQL 1140
1081 DTVDSMIPEVIAKMFMSLFPNVIGACIVILATPAAIIIPPLGLIYFVQRFVASSRQL 1140
1141 KRLIESVSRSPVSHFNETLIGSVIRAFEEQEFTHOSDLKVDENQCAVYPSIVANRWLA 1200
1141 KRLIESVSRSPVSHFNETLIGSVIRAFEEQEFTHOSDLKVDENQCAVYPSIVANRWLA 1200
1201 VRLCEGNCVLEPAAFAVVISRSLAGLVGSVSLQVTVYLVNMLVRRSMETNTIYA 1260
1201 VRLCEGNCVLEPAAFAVVISRSLAGLVGSVSLQVTVYLVNMLVRRSMETNTIYA 1260

Db 1201 VRLCEVNCIVLPAALPAVVISRHSLSAGLVLSYSYSIQVTTYVNLVMSSEMETNIVA 1260
Qy 1261 VERLKEYSETEKEAPWQIOETAPSPSWPOVGRVEFRNYCLRYREDLDFVLHINVTINGG 1320
Db 1261 VERLKEYSETEKEAPWQIOETAPSPSWPOVGRVEFRNYCLRYREDLDFVLHINVTINGG 1320
Qy 1321 EKVGIVGTGAGKSSLTITGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
Db 1321 EKVGIVGTGAGKSSLTITGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
Qy 1381 SSGSRMNLDPSSQVSDSEEWMTSLSLAHKDFVSALPDLDEHCAAGSGNLSVGQROAVCL 1440
Db 1381 SSGSRMNLDPSSQVSDSEEWMTSLSLAHKDFVSALPDLDEHCAAGSGNLSVGQROAVCL 1440
Qy 1441 ARALLRKTIIVLDEATAVLETDLLOSTIRTOFEDCTVLTIAHRLNTIMDYTRIVL 1500
Db 1441 ARALLRKTIIVLDEATAVLETDLLOSTIRTOFEDCTVLTIAHRLNTIMDYTRIVL 1500
Qy 1501 DKGEIOEYGAAPSDLLQORGLFYSMAKDAGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAE 1560
Db 1501 DKGEIOEYGAAPSDLLQORGLFYSMAKDAGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAE 1560
Qy 1561 PRYTRERARFVSKGNCVNAHKNIROGRFLQDVFTTLVDLKMPHLLIFMSFLCSW 1620
Db 1561 PRYTRERARFVSKGNCVNAHKNIROGRFLQDVFTTLVDLKMPHLLIFMSFLCSW 1620
Qy 1621 LLFAMVWMLIAFAHGDLPAGEGTNVCVTSIHSFSSAFLSIEVOVTIGFGGRTVEBCP 1680
Db 1621 LLFAMVWMLIAFAHGDLPAGEGTNVCVTSIHSFSSAFLSIEVOVTIGFGGRTVEBCP 1680
Qy 1681 LAIILIVQNVGLMTNIMIGCTFMKTAQAHRAETLIFSCHAVITTRHGRLCMELRVG 1740
Db 1681 LAIILIVQNVGLMTNIMIGCTFMKTAQAHRAETLIFSCHAVITTRHGRLCMELRVG 1740
Qy 1741 DLKRSMTISATIHMOVAKTTSPEGEVPLHNOVDIPMENGVGNGCIFVAPLIYHVIDS 1800
Db 1741 DLKRSMTISATIHMOVAKTTSPEGEVPLHNOVDIPMENGVGNGCIFVAPLIYHVIDS 1800
Qy 1801 NSPLYDLAPSDLHHQDLEIIVLEGVETTGITTOARTSYLADEILWGORFPIVASED 1860
Db 1801 NSPLYDLAPSDLHHQDLEIIVLEGVETTGITTOARTSYLADEILWGORFPIVASED 1860
Qy 1861 GRYSVDYSKFGNTIKVPTPLCTARQLDEDRS 1891
Db 1861 GRYSVDYSKFGNTIKVPTPLCTARQLDEDRS 1891

RESULT 3
ADY86938
ID ADY86938 standard; protein; 1927 AA.
AC ADY86938;
XX 02-JUN-2005 (first entry)
DE Human MRPI-mouse Klr6.2 fusion protein, SEQ ID NO: 1.
XX Ionophore; biosensor; drug screening; diagnostic;
KW microorganism detection; potassium channel; fusion protein;
KW multidrug resistance protein 1; MRPI; Klr6.2.
XX Homo sapiens.
OS Mus musculus.
OS Chimeric.
XX
FH Key Location/Qualifiers
FH 1..1531
FT Region /note= "Multidrug resistance protein 1 (MRPI)"
FT 1532..1537
FT /note= "Hexaglycine spacer"
FT Region 1538..1927
FT /note= "Klr6.2 protein"
XX

PN US2005063989-A1.
XX
XX 24-MAR-2005.
XX
XX 22-SEP-2003; 2003US-00665283.
XX
XX 22-SEP-2003; 2003US-00665283.
XX
XX (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
XX
XX Detand R, Garcia E, Prost A, Revilloud J, Vivaudou M,
XX WPI; 2005-252611/26.
XX
XX New ion channel hybrid protein, used as electrical sensor for screening
PT an agonist/antagonist of a membrane protein and for detecting a
PT contaminant/pollutant in a sample.
XX
XX Claim 25; SEQ ID NO 1; 78bp; English.
XX
XX The present invention relates to a hybrid protein consisting essentially
CC of the fusion of a membrane protein with an ion channel which is not
CC naturally coupled to the membrane protein. The hybrid protein is used as
CC an electrical sensor for screening of an agonist/antagonist of a membrane
CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multidrug resistance protein 1 (MRPI) and mouse inner rectifier potassium
CC channel Klr6.2 protein.
XX
SQ Sequence 1927 AA:
Query Match 100.0%; Score 9734; DB 9; Length 1927;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALRFGSADGSDPLMDMNTWNTSNPDFYKCFQNTVLAWPCRYLWACPFYFLYLSRH 60
Db 1 MALRFGSADGSDPLMDMNTWNTSNPDFYKCFQNTVLAWPCRYLWACPFYFLYLSRH 60
Qy 61 DRGYQMTPLNKTATGALGFLIMVQWDLFYSFERSRGFLAVPLVSPLLGLITLLA 120
Db 61 DRGYQMTPLNKTATGALGFLIMVQWDLFYSFERSRGFLAVPLVSPLLGLITLLA 120
Qy 121 TFLIQLEERKGVSSGIMLTFWLVALYCALAIIKSKIMTALKEDAQVDLFFRDIFFYVFS 180
Db 121 TFLIQLEERKGVSSGIMLTFWLVALYCALAIIKSKIMTALKEDAQVDLFFRDIFFYVFS 180
Qy 181 LLLIQVLVSCFSDRSPLFSETHDNPCESSASFLSRITTWITGLIVRGYRQPLSGSD 240
Db 181 LLLIQVLVSCFSDRSPLFSETHDNPCESSASFLSRITTWITGLIVRGYRQPLSGSD 240
Qy 241 LMSLNKEDTSQVVPVLYVKNKKCAKTRKQPVVYVSSKDPADPKSSKVDANEVEAL 300
Db 241 LMSLNKEDTSQVVPVLYVKNKKCAKTRKQPVVYVSSKDPADPKSSKVDANEVEAL 300
Qy 301 IVKSPOKEWNPSLFVKLYKTGPFYLSFFPKAIIHDLMMFSGPOILKLLIFVNDTKAPD 360
Db 301 IVKSPOKEWNPSLFVKLYKTGPFYLSFFPKAIIHDLMMFSGPOILKLLIFVNDTKAPD 360
Qy 361 WQGYFYTVLFLVTAQLQTLVLAHQYHICFVSGMRTKTAIVAGVYRKALVITNSARKSSTV 420
Db 361 WQGYFYTVLFLVTAQLQTLVLAHQYHICFVSGMRTKTAIVAGVYRKALVITNSARKSSTV 420
Qy 421 GEIVNLMSVDAQREFDLATYINMTWSAPLOVITLALYLLMLNMGPSVLGAVMVLVNVN 480
Db 421 GEIVNLMSVDAQREFDLATYINMTWSAPLOVITLALYLLMLNMGPSVLGAVMVLVNVN 480
Qy 481 AVMAKKTGYOVAAHKKSKDNRIKLMNEITLNGIKVYKLYAMELAFKDKVLAIRQESLKYLK 540
Db 481 AVMAKKTGYOVAAHKKSKDNRIKLMNEITLNGIKVYKLYAMELAFKDKVLAIRQESLKYLK 540
Qy 541 KSAVLSAVGTFETWCTPFLVALCTFAVYVYTIIDENNIIDAQTAFAVSLALFNILRPLNITLP 600

CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel Kir6.2 protein containing HA epitope peptide.

XX Sequence 1947 AA;

Query Match 99.7%; Score 9704; DB 9; Length 1947;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 20; Gaps 2;

```
QY 1 MALRGFCSADSDPLMDNNTWNTSNPDTKCFQNTVTVWVPCFTLWACFPFPLYSRH 60
DB 1 MALRGFCSADSDPLMDNNTWNTSNPDTKCFQNTVTVWVPCFTLWACFPFPLYSRH 60
QY 61 DRGTYQMTPLNKTALGFLMTICMADLFYSFMRSGIFLAPFLVSPILLGITTLA 120
DB 61 DRGTYQMTPLNKTALGFLMTICMADLFYSFMRSGIFLAPFLVSPILLGITTLA 120
QY 121 TFLIQLEBRKGVSSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVFS 180
DB 121 TFLIQLEBRKGVSSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVFS 180
QY 181 LLLIQVLSCFSDSPFLFSETIHDNPPCPSSASFLSKITFWMTGLIVRGYRQPLEGSD 240
DB 181 LLLIQVLSCFSDSPFLFSETIHDNPPCPSSASFLSKITFWMTGLIVRGYRQPLEGSD 240
QY 241 LMSLNKEDTSQOVVPLVKNMKECAKTRKOPVKVYSSKDPAPKESKVDANEVVAL 300
DB 241 LMSLNKEDTSQOVVPLVKNMKECAKTRKOPVKVYSSKDPAPKESKVDANEVVAL 300
QY 301 IVKSPQKEMNDSLEKVLVKTFGPYLMSPFKAIHIDMMFSGPQILKLILKRVNDTKAP 360
DB 301 IVKSPQKEMNDSLEKVLVKTFGPYLMSPFKAIHIDMMFSGPQILKLILKRVNDTKAP 360
QY 361 WQGFYTYVLLVNTACLQTLVHQYHICFVSGMRKTAIVIGAVRKALVITNSARKSTV 420
DB 361 WQGFYTYVLLVNTACLQTLVHQYHICFVSGMRKTAIVIGAVRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDARFMDLTYINMISAPLOVYIALYLLMLNGSPVLGAVVMVLMVNVN 480
DB 421 GEIVNLMSVDARFMDLTYINMISAPLOVYIALYLLMLNGSPVLGAVVMVLMVNVN 480
QY 481 AVMAKTKTYOVANHKSKDNRIKLMNEILNGIKVLKLYAMBLAFKQVLAIROBELKVLK 540
DB 481 AVMAKTKTYOVANHKSKDNRIKLMNEILNGIKVLKLYAMBLAFKQVLAIROBELKVLK 540
QY 541 KSAIYLSAVGTWCTPFLVALCTPAVYVTTIDENNIIDAQTAFAVSLAFNILRPFLNLP 600
DB 541 KSAIYLSAVGTWCTPFLVALCTPAVYVTTIDENNIIDAQTAFAVSLAFNILRPFLNLP 600
QY 601 MVISSIVASVSLKRLRIFLSHBELEPDSIERRPVKDGGTNSIVRNATFTWASDPPT 660
DB 601 MVISSIVASVSLKRLRIFLSHBELEPDSIERRPVKDGGTNSIVRNATFTWASDPPT 660
QY 661 LMGITFSPREGALVAVVGQVGKSSLSALLAEMDKVGHVAKGSVAAYVPOQAMIOND 720
DB 661 LMGITFSPREGALVAVVGQVGKSSLSALLAEMDKVGHVAKGSVAAYVPOQAMIOND 720
QY 721 SLRENILFGCQLEBPYRSVIOACALLPDLEILPSGDRTEIGEKNVNSGGQKQVSLAR 780
DB 721 SLRENILFGCQLEBPYRSVIOACALLPDLEILPSGDRTEIGEKNVNSGGQKQVSLAR 780
QY 781 AVYNSADIYLPDDPLSAVDAAVGHKHFENNVGPKMLKNKRIILVTHMSMYLPQVDVITV 840
DB 781 AVYNSADIYLPDDPLSAVDAAVGHKHFENNVGPKMLKNKRIILVTHMSMYLPQVDVITV 840
QY 841 MSGGKISEMGSYOEILLADGAFAEFLRTYASTEQDQAEENGVTGSGPGEAKQEMNGM 900
DB 841 MSGGKISEMGSYOEILLADGAFAEFLRTYASTEQDQAEENGVTGSGPGEAKQEMNGM 900
QY 901 LVTDAGKOLQROLSSSSSYSGDISRHNSTAELOKAEAKKEETWKLMEADKQOTGOVKL 960
DB 901 LVTDAGKOLQROLSSSSSYSGDISRHNSTAELOKAEAKKEETWKLMEADKQOTGOVKL 960
```

```
DB 901 LVTDAGKOLQROLSSSSSYSGDISRHNSTAELOKAEAKKEETWKLMEADKQOTGOVKL 960
QY 961 SVYDWYKAIQGLFISFLSIFLFCMCHVSALASNWYLSLWTDDPVINGTOEHTKVRLSYVG 1020
DB 961 SVYDWYKAIQGLFISFLSIFLFCMCHVSALASNWYLSLWTDDPVINGTOEHTKVRLSYVG 1020
QY 1021 ALGISQGIAVFGYSMAVSIQGIILASRCLHYDLHLSIIRSPMSFPERTPSGULVNRFSKEL 1080
DB 1021 ALGISQGIAVFGYSMAVSIQGIILASRCLHYDLHLSIIRSPMSFPERTPSGULVNRFSKEL 1080
QY 1081 DTVDMSIPEVTKMMSGLPNVIGACIVILLATPIAIIIPPLGLIYFFVORFYVASSROL 1140
DB 1081 DTVDMSIPEVTKMMSGLPNVIGACIVILLATPIAIIIPPLGLIYFFVORFYVASSROL 1140
QY 1141 KRLESVRSPPVSHFNSTLLGVSYIRAFEEQERFIHOSDLKVDENQAKAYPSIVANRWLA 1200
DB 1141 KRLESVRSPPVSHFNSTLLGVSYIRAFEEQERFIHOSDLKVDENQAKAYPSIVANRWLA 1200
QY 1201 VRLCEVNCIYLPALPAVVISRHSLSAGLVLSYSISLOTTYINMLVRRSSEMETNIVA 1260
DB 1201 VRLCEVNCIYLPALPAVVISRHSLSAGLVLSYSISLOTTYINMLVRRSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMGOIOETAPPSWPOVGRVFERNYCLRYREDLPFLRHIVTTINGG 1320
DB 1261 VERLKEYSETEKEAPMGOIOETAPPSWPOVGRVFERNYCLRYREDLPFLRHIVTTINGG 1320
QY 1321 EKVGIIVRTGAKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPDPVLF 1380
DB 1321 EKVGIIVRTGAKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPDPVLF 1380
QY 1381 SGIIRPMULDPPSOYSDBEWTSLELAHKDFVSLPKLDBHECAEGENLSVGOQOLVCL 1440
DB 1381 SGIIRPMULDPPSOYSDBEWTSLELAHKDFVSLPKLDBHECAEGENLSVGOQOLVCL 1440
QY 1441 ARALLRKTKLIVDEAPRAVDLEETDILQSTIRPOFECTVLTAFHRLNTMDTRVYVL 1500
DB 1441 ARALLRKTKLIVDEAPRAVDLEETDILQSTIRPOFECTVLTAFHRLNTMDTRVYVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMAXDAGLVGGGGGGLSRKGIIPBEVYVLTSLAEPPAE 1560
DB 1501 DKGEIOEYGAPSDLLQORGLFYSMAXDAGLVGGGGGGLSRKGIIPBEVYVLTSLAEPPAE 1560
QY 1561 PRYTRERRARFVSKKGCNVAAHKNIRBQGRFLQDVFTTVLDLKWPHTLIIFWMSFLCSW 1620
DB 1561 PRYTRERRARFVSKKGCNVAAHKNIRBQGRFLQDVFTTVLDLKWPHTLIIFWMSFLCSW 1620
QY 1621 LLFPMWMLTAFAG-----DLAP-----GGRNVPCVTSIHSFSAFLF 1660
DB 1621 LLFPMWMLTAFAGHDLYAYMEKGITDLAPYPVDPYAGGNTVPCVTSIHSFSAFLF 1660
QY 1661 SIEVQVTIGFGRAWTEECPLAILILIVONIVGMINAIMLGCIFMKTAAQHRAEETLIF 1720
DB 1661 SIEVQVTIGFGRAWTEECPLAILILIVONIVGMINAIMLGCIFMKTAAQHRAEETLIF 1720
QY 1721 SKHAVITLRHGRCLCFMLRVGDLRKSMIISATIHMQVVRKTTSPGEEVVPDLHQVDIPMENG 1780
DB 1721 SKHAVITLRHGRCLCFMLRVGDLRKSMIISATIHMQVVRKTTSPGEEVVPDLHQVDIPMENG 1780
QY 1781 VGGNGIFLVAAPLIYHYIDSNSPLYDLAPSLHHHODEIIVILEGVETGITTOARTS 1840
DB 1781 VGGNGIFLVAAPLIYHYIDSNSPLYDLAPSLHHHODEIIVILEGVETGITTOARTS 1840
QY 1841 YLADEIILMGQRFPIVAEEDGRYSVDYSKFGNTIKVPTPLCTARQOLDEDRS 1891
DB 1841 YLADEIILMGQRFPIVAEEDGRYSVDYSKFGNTIKVPTPLCTARQOLDEDRS 1891
RESULT 5
AAW57486
ID AAW57486 standard; protein; 1531 AA.
XX
AC AAW57486;
XX
```

DT 14-AUG-1998 (first entry)
XX Human MRP variant ltpgpa (lei/pgpa).
XX Multidrug resistance-associated protein; MRP; tumour; human; variant;
KM multidrug resistance; MDR; leismania P-glycoprotein; ltpgpa; lei/pgpa.
XX
OS Homo sapiens.
FH
FH Key Location/Qualifiers
FT Misc-difference 685
FT /label= L685S
FT /note= "wild-type leu is replaced by ser"
FT Misc-difference 1282
FT /label= R1282A
FT /note= "wild-type Arg is replaced by Ala"
XX US5766880-A.
XX 16-JUN-1998.
XX 05-JUN-1995; 95US-00463092.
XX 27-OCT-1992; 92US-00966923.
XX 08-MAR-1993; 93US-00029340.
XX 26-OCT-1993; 93US-00141893.
XX 20-MAR-1995; 95US-00407207.
XX (TOOH) UNIV QUEBENS KINGSTON.
XX
XX Cole SP, Deeley RG;
XX
XX WPI: 1998-361687/31.
XX N-PSDB; AAV31498.
XX
XX DNA encoding protein associated with multi-drug resistance - useful for
PT as probe for identifying multi-drug resistant tumour cells.
XX
XX
XX Claim 1; Col 67-78; 82pp; English.
XX
XX This represents a variant of the human multidrug resistance-associated
CC protein (MRP). This natural variant is a leismania P-glycoprotein related
CC molecule ltpgpa (lei/pgpa). The human and murine MRP nucleic acid
CC molecules can be used as probes for identifying multidrug resistant
CC tumour cells by hybridisation to mRNA from tumour cells. The antisense
CC nucleic acid can be used to reverse multidrug resistance (MDR). A
CC recombinant expression vector containing the MRP nucleic acid molecules
CC operatively linked to at least one regulatory sequence can be used to
CC transform a host cell to produce a recombinant MDR-associated protein
XX
SQ Sequence 1531 AA:
Query Match 80.7%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 241 LMSLNKEDTSEQVAVLVKQWKKECAKTRKQPVKVVYSSKDPAQPKSSSKVDANEVEAL 300
QY 301 IVSPQKKNPISLFKVLKTPGPFYFLMSFFPKAIDLMFSGPOILKLLIKFVNDTAPD 360
DB 301 IVSPQKKNPISLFKVLKTPGPFYFLMSFFPKAIDLMFSGPOILKLLIKFVNDTAPD 360
QY 361 MGGFYFVLLPVTAQCTQTLVHOVPHI CFSGMR IKTAVIGAVRKALVTNSARKSTV 420
DB 361 MGGFYFVLLPVTAQCTQTLVHOVPHI CFSGMR IKTAVIGAVRKALVTNSARKSTV 420
QY 421 GEIVNLMSVDAQREMDATYINMISAPLOVITLALYLMINIGPSVLAGVAVMLAMPVN 480
DB 421 GEIVNLMSVDAQREMDATYINMISAPLOVITLALYLMINIGPSVLAGVAVMLAMPVN 480
QY 481 AVNARKTXYTVQAHMKSNDRIKLMNEILINGIKLYAMELAFKDKVLAIROBELKVLK 540
DB 481 AVNARKTXYTVQAHMKSNDRIKLMNEILINGIKLYAMELAFKDKVLAIROBELKVLK 540
QY 541 KSAVLAVGTFPTWCPEFVALCTFAVYVITDENNILDQAQAFVSLAFNLRPEPLILP 600
DB 541 KSAVLAVGTFPTWCPEFVALCTFAVYVITDENNILDQAQAFVSLAFNLRPEPLILP 600
QY 601 MVISIIYQASVSLKRLRIFLSHEELBPSIERRPVKDGGTNSITVENATFTMARSPPT 660
DB 601 MVISIIYQASVSLKRLRIFLSHEELBPSIERRPVKDGGTNSITVENATFTMARSPPT 660
QY 661 LMGITFSIPGALVAVVGVCGCKSSLSALLAEMDRVGHVAIKGSVAVYPOQAWI QND 720
DB 661 LMGITFSIPGALVAVVGVCGCKSSLSALLAEMDRVGHVAIKGSVAVYPOQAWI QND 720
QY 721 SLRENILFGQLEBPYRSYIOACALLPDLIELPBGRTETGEGVNLSSGQKRVSLAR 780
DB 721 SLRENILFGQLEBPYRSYIOACALLPDLIELPBGRTETGEGVNLSSGQKRVSLAR 780
QY 781 AVYSNADITYLPDDPLSAVDAAVGHIFENYIGPKMKAKTRILVITMSMSTLPQVDYIIV 840
DB 781 AVYSNADITYLPDDPLSAVDAAVGHIFENYIGPKMKAKTRILVITMSMSTLPQVDYIIV 840
QY 841 MSGGKISEMGSYOELARDAFAEFLRTVASTBOEDABENGVTGVSFGPKKAKOMENGM 900
DB 841 MSGGKISEMGSYOELARDAFAEFLRTVASTBOEDABENGVTGVSFGPKKAKOMENGM 900
QY 901 LVYDSACKQIQORQLSSSSSYSGDISRRHNSTAEIQKAEAKETWKLMEADKAQTGVKL 960
DB 901 LVYDSACKQIQORQLSSSSSYSGDISRRHNSTAEIQKAEAKETWKLMEADKAQTGVKL 960
QY 961 SVYWDYKAKIGLFIISFSLIFLFCMNHVSALASNYWLSLMTDDPIVNGTOSHTRVLSVYG 1020
DB 961 SVYWDYKAKIGLFIISFSLIFLFCMNHVSALASNYWLSLMTDDPIVNGTOSHTRVLSVYG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLHVDLHLSILRSPMSFEERTPSGNLVNRPSEKL 1080
DB 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLHVDLHLSILRSPMSFEERTPSGNLVNRPSEKL 1080
QY 1081 DTYDSMIPETIYKMPGSLFNVIACVITLATPAAIIPPLGLIYFFVORFYVASSROL 1140
DB 1081 DTYDSMIPETIYKMPGSLFNVIACVITLATPAAIIPPLGLIYFFVORFYVASSROL 1140
QY 1141 KRLSVSRSPVYSHFNETLIGSVITRAFEBQERFIHOSDLKVDENQAKAYPSIVANRWLA 1200
DB 1141 KRLSVSRSPVYSHFNETLIGSVITRAFEBQERFIHOSDLKVDENQAKAYPSIVANRWLA 1200
QY 1201 VRLCEVGNCTVLPALFAVISRSLSAGVLGSVSYSLQVTTYIAMLVRMSSEMETIVA 1260
DB 1201 VRLCEVGNCTVLPALFAVISRSLSAGVLGSVSYSLQVTTYIAMLVRMSSEMETIVA 1260
QY 1261 VERLKEYSERTEKAPMOIQETAPSSWPQGRVGFENYCLIRYEDDLPVLRHINVTINGG 1320
DB 1261 VERLKEYSERTEKAPMOIQETAPSSWPQGRVGFENYCLIRYEDDLPVLRHINVTINGG 1320
QY 1321 EKVGIAGRTAGAGSSLTGLFRINESAEGRIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIAGRTAGAGSSLTGLFRINESAEGRIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380

Db 1321 EKVGIVGRTGAGKSLTGLFRINESAGEI110GINIAKIGLHDLRPFKTIIPDPVL 1380
Qy 1381 SGLSRMLDPSQXSDEEWTSLSLAHKDPVSAIPDLDEHCAEGENLSVGOROLVCL 1440
Db 1381 SGLSRMLDPSQXSDEEWTSLSLAHKDPVSAIPDLDEHCAEGENLSVGOROLVCL 1440
Qy 1441 ARALLRKTKIIVLDEATPAVDLETFDDLIQSTIRQFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Db 1441 ARALLRKTKIIVLDEATPAVDLETFDDLIQSTIRQFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Qy 1501 DKGEIOERYGAPSDLLQORGLFYSMAKDAGLV 1531
Db 1501 DKGEIOERYGAPSDLLQORGLFYSMAKDAGLV 1531
RESULT 6
AAW74471
ID AAW74471 standard; protein, 1531 AA.
XX
AC AAW74471;
XX
DT 18-MAY-1999 (first entry)
XX
DS Human multidrug resistance-associated protein variant.
XX
KM Multidrug resistance-associated protein; MDR; human; diagnosis;
XX MDR tumour cell identification; cancer therapy.
XX
OS Homo sapiens.
XX
PN US5882875-A.
XX
PD 16-MAR-1999.
XX
PE 05-JUN-1995; 95US-00462109.
XX
PR 27-OCT-1992; 92US-00966923.
PR 08-MAR-1993; 93US-00029340.
PR 26-OCT-1993; 93US-00141893.
PR 20-MAR-1995; 95US-00407207.
XX
PA (TOOH) UNIV QUEBENS KINGSTON.
PI Cole SPC, Deeley RC;
XX
DR WPI: 1999-214061/18.
DR N-PSDB; AAX21977.
XX
PT Identifying a multidrug resistant tumor cell by contacting the cell with
PT an antibody/antigen-binding fragment - which binds to an expressed
PT protein encoded by multidrug resistance-associated protein (MRP) nucleic
PT acid.
XX
PS Claim 3; Col 69-80; 80pp; English.
XX
CC This sequence is the human multidrug resistance-associated (MDR) protein.
CC The invention relates to a method for identifying a multidrug resistant
CC (MDR) tumour cell. Compositions and methods utilising the MDR proteins
CC can be used to treat patients with tumours displaying multidrug
CC resistance, particularly those displaying resistance to anthracyclines,
CC epipodophylotoxins, vinca alkaloids, and hydrophobic drugs. The methods
CC for inhibiting/killing a MDR tumour cell can be useful for treating
CC breast cancer, leukemias, fibrosarcomas, cervical cancer, gliomas
CC thymomas, neuroblastomas and lung cancer. The MDR DNA sequences when
CC labeled are useful as molecular probes for diagnosing multidrug
CC resistance of a tumour (using cells from a tumour biopsy) and for
CC designing ribozymes which are capable of cleaving a single-stranded
CC nucleic acid encoding a protein having MRP activity. Recombinant
CC expression vectors containing human MDR coding sequences can be
CC transfected into a drug sensitive cell line to produce a protein in the
CC cell which confers MDR, protecting non-resistant non-tumour cells from
CC the effects of chemotherapeutics has major clinical importance. Cells
CC transformed with the MDR coding sequences are useful for testing

CC potential therapeutic agents for their effectiveness against MDR cells
CC and for identifying chemosensitizers of a therapeutic agent
XX
SQ Sequence 1531 AA;
Query Match 80.7%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALRGFCSAGSDPLMDMNTVMTNTPNDPFCQONTVLVWVPCFYLAACPFYFLYLSRH 60
Db 1 MALRGFCSAGSDPLMDMNTVMTNTPNDPFCQONTVLVWVPCFYLAACPFYFLYLSRH 60
Qy 61 DRGIQMTPLNKTKTALGFLIMIVCMADLFYSFWRSGIPLAEVPLVSPTLGITTLLA 120
Db 61 DRGIQMTPLNKTKTALGFLIMIVCMADLFYSFWRSGIPLAEVPLVSPTLGITTLLA 120
Qy 121 TFLIQLEBRKGVSSGIMLTFWLVALYCALAILRSKIMTALKEDAOVDLPFDITFYFFS 180
Db 121 TFLIQLEBRKGVSSGIMLTFWLVALYCALAILRSKIMTALKEDAOVDLPFDITFYFFS 180
Qy 181 LLLIQVLVSCPSDRSPLFSETIHDNPPCPSSASFLSRITPFWITGLIVRGYRQPLEGSD 240
Db 181 LLLIQVLVSCPSDRSPLFSETIHDNPPCPSSASFLSRITPFWITGLIVRGYRQPLEGSD 240
Qy 241 LMSLNKEDTSEQVVPVLVKNWKKCECAKTRKQPVKVVYSSKDPAPKSSSKVDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLVKNWKKCECAKTRKQPVKVVYSSKDPAPKSSSKVDANEVEAL 300
Qy 301 IVKSPQKBNPDSLFRVLYKTFGPFLMSFFPKA1HDLMSRGPOTLKLLIFVNDTAPD 360
Db 301 IVKSPQKBNPDSLFRVLYKTFGPFLMSFFPKA1HDLMSRGPOTLKLLIFVNDTAPD 360
Qy 361 WQGFYTVLTFVTACLTQTLVHOYFHI CFSGMS IKTAVI GAVYRKALVITNSARKSTV 420
Db 361 WQGFYTVLTFVTACLTQTLVHOYFHI CFSGMS IKTAVI GAVYRKALVITNSARKSTV 420
Qy 421 GEIVNLMSVDAQRPMDLATYINMISAPLQVITLALYLLMNLGPSVLAVAVWVLMVPVN 480
Db 421 GEIVNLMSVDAQRPMDLATYINMISAPLQVITLALYLLMNLGPSVLAVAVWVLMVPVN 480
Qy 481 AVMAKTKTYOVAMKSKDNRIKLMNEILINGIKVLKYAMELAKDKVLA1ROBELKYLK 540
Db 481 AVMAKTKTYOVAMKSKDNRIKLMNEILINGIKVLKYAMELAKDKVLA1ROBELKYLK 540
Qy 541 KSAVLSAVGFTWCTPFIVALCTFAVYVITDENNNIIDAQTAFLALFNILRPPLNLP 600
Db 541 KSAVLSAVGFTWCTPFIVALCTFAVYVITDENNNIIDAQTAFLALFNILRPPLNLP 600
Qy 601 MVISIVQASVSLKRLIFLSHEELPDSIERRPVKDGGTNSITVNAATPTWARSDEPT 660
Db 601 MVISIVQASVSLKRLIFLSHEELPDSIERRPVKDGGTNSITVNAATPTWARSDEPT 660
Qy 661 LNTGTFSPBEGALVAVVQVCGKSSLLSALLAMDEKGVHAIKGSVAVYPOQAMIOND 720
Db 661 LNTGTFSPBEGALVAVVQVCGKSSLLSALLAMDEKGVHAIKGSVAVYPOQAMIOND 720
Qy 721 SLRENILFGCLLEPPYRSVYQACALLPDLEILSISGRTETGEGVNLSSGQKRVSLAR 780
Db 721 SLRENILFGCLLEPPYRSVYQACALLPDLEILSISGRTETGEGVNLSSGQKRVSLAR 780
Qy 781 AVYNSADLYLPDDPLSAVDAAVGHIFENYVGPKGMLKNKRIIVTHSMYSYLPQDVYIV 840
Db 781 AVYNSADLYLPDDPLSAVDAAVGHIFENYVGPKGMLKNKRIIVTHSMYSYLPQDVYIV 840
Qy 841 MSGKISMSGSYOBLARDAFAEFLRTYASTEOBDAEENGVTGVSQPGKBAKOMENG 900
Db 841 MSGKISMSGSYOBLARDAFAEFLRTYASTEOBDAEENGVTGVSQPGKBAKOMENG 900
Qy 901 LVTSAGKOLOROUSSSSSYGDISRHNTGTAELQXAKKEETWKLMEADKAOYGQYKL 960
Db 901 LVTSAGKOLOROUSSSSSYGDISRHNTGTAELQXAKKEETWKLMEADKAOYGQYKL 960

QY 961 SVYDYMKAIGLFLISFLISFLMCMNHVSALASNYWLSMTDDPIYNGTOEHTKYRLSVYG 1020
DB 961 SVYDYMKAIGLFLISFLISFLMCMNHVSALASNYWLSMTDDPIYNGTOEHTKYRLSVYG 1020
QY 1021 ALGISOGIAVGVMAVSIIGIILASRCLHVDLHSLIRS PMSFFERTPSGNLVNRFSEKL 1080
DB 1021 ALGISOGIAVGVMAVSIIGIILASRCLHVDLHSLIRS PMSFFERTPSGNLVNRFSEKL 1080
QY 1081 DTYSMSIPEVTKMFMGSLFNVIIGACIVTLATPIAIIIPPLGLIYFVQGFYVASSKQL 1140
DB 1081 DTYSMSIPEVTKMFMGSLFNVIIGACIVTLATPIAIIIPPLGLIYFVQGFYVASSKQL 1140
QY 1141 KRLESVSRSPYSHFNELLGVSVIRAFEEQERFIHOSDLKYDENOKAYYSIVANRWLA 1200
DB 1141 KRLESVSRSPYSHFNELLGVSVIRAFEEQERFIHOSDLKYDENOKAYYSIVANRWLA 1200
QY 1201 VRLCEVGNCLVLPALPAVISRHSLSAGLVLSVYSIQVTTYNLWLVMSSEMETNIVA 1260
DB 1201 VRLCEVGNCLVLPALPAVISRHSLSAGLVLSVYSIQVTTYNLWLVMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMOIQETAPSSWPQVGRVFRNCLAYREDLDPVLAHINVTINGG 1320
DB 1261 VERLKEYSETEKEAPMOIQETAPSSWPQVGRVFRNCLAYREDLDPVLAHINVTINGG 1320
QY 1321 EKVGIVERTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRPKITIIPQDPVLF 1380
DB 1321 EKVGIVERTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRPKITIIPQDPVLF 1380
QY 1381 SGIIRNMLDPPSQYSDEEVTSLSLAHLKDFVSALPDKLDHECAEGENLSVQROQLVCL 1440
DB 1381 SGIIRNMLDPPSQYSDEEVTSLSLAHLKDFVSALPDKLDHECAEGENLSVQROQLVCL 1440
QY 1441 ARALLRTKILVDEATPAVDLETFDDLIQSTIRIQFEDCTVLTAAHLANTIMDYTRYIVL 1500
DB 1441 ARALLRTKILVDEATPAVDLETFDDLIQSTIRIQFEDCTVLTAAHLANTIMDYTRYIVL 1500
QY 1501 DKGEIOEGAPSDLLQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEGAPSDLLQORGLFYSMAKDAGLV 1531

RESULT 7
AAM99894
ID AAM99894 standard; protein; 1531 AA.
XX
AC AAM99894;
XX
DT 10-JUN-1999 (first entry)
XX
DE Human multidrug resistance-associated protein natural variant.
XX
KW Human; multidrug resistance-associated protein; MRP; cytotoxic drug;
KM cancer; chemotherapy.
XX
OS Homo sapiens.
XX
PN US5891724-A.
XX
PD 06-APR-1999.
XX
PE 05-JUN-1995; 95US-00460907.
XX
PR 27-OCT-1992; 92US-00966923.
PR 08-MAR-1993; 93US-00029340.
PR 26-OCT-1993; 93US-00141893.
PR 20-MAR-1995; 95US-00407207.
XX
PA (TOOH) UNIV QUEBENS KINGSTON.
XX
PI Cole SPC, Deeley RG;
XX
XX WPI; 1999-253868/21.
DR N-PSDB; AAX19818.

XX
PT Protecting mammalian cells against cytotoxic drugs.
XX
PS Claim 2; Col 79-86; 82pp; English.
XX
CC The present sequence represents a human multidrug resistance-associated
CC protein (MRP). The present invention also describes a method for
CC protecting a mammalian cell against the cytotoxicity of anthracyclines,
CC epipodophyllotoxins and vinca alkaloids (A) by introducing into it a
CC nucleic acid (I) that hybridizes under stringent conditions to a nucleic
CC acid (II) that encodes an MRP protein (III). Introduction of (I) protects
CC cells against cytotoxic effects of (A), particularly to protect normal
CC cells against (A) being used for treatment of cancers. Cells transformed
CC with (I) can be used to screen for agents that affect multidrug
CC resistance or are directly toxic to multidrug resistant cells, i.e.
CC potential therapeutics for multidrug-resistant cancers. Confering
CC resistance to normal cells should allow an increase in the dose of (A)
CC that can be administered safely
XX

SO Sequence 1531 AA;

Query Match 80.7%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTNTNSNDFPKCFQNTLVWVPCYILMACFPFYLYSRH 60
DB 1 MALRGFCSADGSDPLMDMNTNTNSNDFPKCFQNTLVWVPCYILMACFPFYLYSRH 60
QY 61 DRGVIQMTPLNKTKTALGFLMTIVCMADLFYSFMRSGIIFLAPVFLVSLGITTLLA 120
DB 61 DRGVIQMTPLNKTKTALGFLMTIVCMADLFYSFMRSGIIFLAPVFLVSLGITTLLA 120
QY 121 TFLIQERRKGVSSGIMLTFWLVALYCALAIIIRSKIMTALKDAQVDLFRDITFYVYS 180
DB 121 TFLIQERRKGVSSGIMLTFWLVALYCALAIIIRSKIMTALKDAQVDLFRDITFYVYS 180
QY 181 LLLIQVLSCFSDRSPFSETHIDPNPCPSSASFSLRIFPMWITGLIYNGROPLEGSD 240
DB 181 LLLIQVLSCFSDRSPFSETHIDPNPCPSSASFSLRIFPMWITGLIYNGROPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVVKMKKECAKTRQOPKVVYSSDPAOPKSSKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVVKMKKECAKTRQOPKVVYSSDPAOPKSSKVDANEVEAL 300
QY 301 IVKSPQKWNPSLFKVLKTFPGYFLMSFFFKAIHDLMSGPQILKLIKFVNDTRAPD 360
DB 301 IVKSPQKWNPSLFKVLKTFPGYFLMSFFFKAIHDLMSGPQILKLIKFVNDTRAPD 360
QY 361 WQGYFTVLLFVNAQLOTLVLAHQYFHCFSNGRIKTAIVGAYRRKALVITNSARKSSTV 420
DB 361 WQGYFTVLLFVNAQLOTLVLAHQYFHCFSNGRIKTAIVGAYRRKALVITNSARKSSTV 420
QY 421 GEIVNMSVDAQRFMDLATYINNIWSAPLOVITLALYLMNLPVSLAGVAVWLVMPVN 480
DB 421 GEIVNMSVDAQRFMDLATYINNIWSAPLOVITLALYLMNLPVSLAGVAVWLVMPVN 480
QY 481 AVNAKTKTYQVAHMSKDNRIKLMEIILNGIKVLKLYAMELAFKQVLAIRQEBELVKL 540
DB 481 AVNAKTKTYQVAHMSKDNRIKLMEIILNGIKVLKLYAMELAFKQVLAIRQEBELVKL 540
QY 541 KSAVLNAVGTFTVCTPFLVALCTFAVYVITIDENNILDAQTAVSLAFNIIARPLNLP 600
DB 541 KSAVLNAVGTFTVCTPFLVALCTFAVYVITIDENNILDAQTAVSLAFNIIARPLNLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHELEPDSIERPVDGGGTSNITRNATFTTARSDEPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHELEPDSIERPVDGGGTSNITRNATFTTARSDEPT 660
QY 661 LNGITFSIPGALVAVVGQVCGKSSLSALLAEKMDKVBGHVALIKSVAVVPOOAMIOND 720
DB 661 LNGITFSIPGALVAVVGQVCGKSSLSALLAEKMDKVBGHVALIKSVAVVPOOAMIOND 720

```

Qy 721 SLRENIIFGCGLEPPYRSVIQACALLPDLIELPSGDRTEIGEGKVNISGGQKORVSLAR 780
Db 721 SLRENIIFGCGLEPPYRSVIQACALLPDLIELPSGDRTEIGEGKVNISGGQKORVSLAR 780
Qy 781 AAVENADIYLPDDLSAVDAHVGHIPENVIQPGKMLKNKRIILVTHSMSTLPQVDVIIV 840
Db 781 AAVENADIYLPDDLSAVDAHVGHIPENVIQPGKMLKNKRIILVTHSMSTLPQVDVIIV 840
Qy 841 MSGGKISEMGSYQELIARDAFAEFLRTYASTEOBDAEENGVTGVSPGKEAKOMENGM 900
Db 841 MSGGKISEMGSYQELIARDAFAEFLRTYASTEOBDAEENGVTGVSPGKEAKOMENGM 900
Qy 901 LVTDASAGKOLQROLSSSSSYSGDISRHHNSTAELOKAKAKKEETWKLMEADKAQTQYKL 960
Db 901 LVTDASAGKOLQROLSSSSSYSGDISRHHNSTAELOKAKAKKEETWKLMEADKAQTQYKL 960
Qy 961 SVYWDYMKATIGLFTSFLSILFMCNHSALASNYWLMTDPPVINGOETHKXRLSYVG 1020
Db 961 SVYWDYMKATIGLFTSFLSILFMCNHSALASNYWLMTDPPVINGOETHKXRLSYVG 1020
Qy 1021 ALGISQGIABVGYGMAVSIIGILASRCLHVDLHSHILSPMSFEERTSPGNLVNRFKEL 1080
Db 1021 ALGISQGIABVGYGMAVSIIGILASRCLHVDLHSHILSPMSFEERTSPGNLVNRFKEL 1080
Qy 1081 DTVDSMIPEVIKMGSLFNVIAGCTIVILLATPIAIIIPPLGLIFFVQRPVYASSRQL 1140
Db 1081 DTVDSMIPEVIKMGSLFNVIAGCTIVILLATPIAIIIPPLGLIFFVQRPVYASSRQL 1140
Qy 1141 KRLSVSRPYVSHFNFTLLGVSVTRAFEBQERFHQSDLVNDENQKAYYSIVANRWLA 1200
Db 1141 KRLSVSRPYVSHFNFTLLGVSVTRAFEBQERFHQSDLVNDENQKAYYSIVANRWLA 1200
Qy 1201 VRLCEVGNICIVLFAVIRSRHSAGLVGSYSIQVTTYLNMLVRRSMSEMETNIVA 1260
Db 1201 VRLCEVGNICIVLFAVIRSRHSAGLVGSYSIQVTTYLNMLVRRSMSEMETNIVA 1260
Qy 1261 VERLKEVSETEKAPWQIQETAPPSWPQVGRVEFRNYCLARRREDDLVFLHINVTINGG 1320
Db 1261 VERLKEVSETEKAPWQIQETAPPSWPQVGRVEFRNYCLARRREDDLVFLHINVTINGG 1320
Qy 1321 EKVGIUGRTGAGKSLTGLFRINSAGEIITIOSINAKIGHLRKRTIIPDDPLF 1380
Db 1321 EKVGIUGRTGAGKSLTGLFRINSAGEIITIOSINAKIGHLRKRTIIPDDPLF 1380
Qy 1381 SGLSRMNLDPSPQYSDEEWMSTLELAHLKDFVSALPDKLDHECAGGENTLVSQORQLVCL 1440
Db 1381 SGLSRMNLDPSPQYSDEEWMSTLELAHLKDFVSALPDKLDHECAGGENTLVSQORQLVCL 1440
Qy 1441 ARALLRKTKIILVDEATAVADLETDLLIQSTIRTFEDCTVLTIAHRLNTIMDYTRIVL 1500
Db 1441 ARALLRKTKIILVDEATAVADLETDLLIQSTIRTFEDCTVLTIAHRLNTIMDYTRIVL 1500
Qy 1501 DKGEIQEYGARSDLLQORGLFRYSMAKXGLV 1531
Db 1501 DKGEIQEYGARSDLLQORGLFRYSMAKXGLV 1531

```

RESULT 8

AAVS5799
ID AAVS5799 standard, protein, 1531 AA.

AAVS5799;
AC AAVS5799;

DT 28-FEB-2000 (first entry)

DB Human multidrug resistance-associated protein (MRP) variant.

KW Chemosensitizer; multidrug resistance-associated protein; MRP; human;
KW therapeutic agent; P-glycoprotein-mediated multidrug resistance; lung;
KW cancer; variant.

OS Homo sapiens.

XX

```

FH Key Location/Qualifiers
FT Misc-difference 685
FT /label= L685S
FT /note= "wild-type Leu is replaced by Ser"
FT Misc-difference 1282
FT /label= R1282A
FT /note= "wild-type Arg is replaced by Ala"
PN US6001563-A.
XX
XX 14-DEC-1999.
PD
XX
XX 05-JUN-1995; 95US-00463179.
PE
XX
XX 27-OCT-1992; 92US-00966923.
PR 08-MAR-1993; 93US-00029340.
PR 26-OCT-1993; 93US-00141893.
PR 20-MAR-1995; 95US-00407207.
XX
XX (TOOH ) UNIV QUEBENS KINGSTON.
PA
XX
XX Cole SP, Deeley RG;
PI WPI, 2000-061877/05.
XX
XX N-PSDB; AAZ39556.
DR
DR Identification of chemosensitizers useful for treating cancer, using
PT nucleic acids encoding multidrug resistance-associated protein.
XX
XX Claim 3; Col 69-80; 77pp; English.
PS
XX
XX The invention provides a method for identifying a substance which is a
CC chemosensitizer that comprises, contacting a cell transfected with
CC nucleic acid encoding multidrug resistance-associated protein (MRP) with
CC a therapeutic agent in vitro. The method is useful for identifying
CC chemosensitizers which may then be used to treat cancer (especially lung
CC cancer). The method allows the identification of chemosensitizers which
CC do not reverse P-glycoprotein-mediated multidrug resistance. The present
CC sequence represents a human MRP variant
XX
SQ Sequence 1531 AA:
Query Match 80.7%; Score 7860; DB 3; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALRGFSGADSDPLMDMNTVMTNSNPDFTKCFQNTVLAWVPCFYLMACFPFYFLYLSRH 60
Db 1 MALRGFSGADSDPLMDMNTVMTNSNPDFTKCFQNTVLAWVPCFYLMACFPFYFLYLSRH 60
Qy 61 DRGYIQMTPLNKTKTALGFLIMVCMADLFYSPWERSRGFLIAPVFLVSPTLIGITLLA 120
Db 61 DRGYIQMTPLNKTKTALGFLIMVCMADLFYSPWERSRGFLIAPVFLVSPTLIGITLLA 120
Qy 61 TFLIQLEBRKGVQSSGIMLTFWLVAVCALAIIRSKIMTALKEDAOYDLFDDITFYVYS 180
Db 61 TFLIQLEBRKGVQSSGIMLTFWLVAVCALAIIRSKIMTALKEDAOYDLFDDITFYVYS 180
Qy 121 TFLIQLEBRKGVQSSGIMLTFWLVAVCALAIIRSKIMTALKEDAOYDLFDDITFYVYS 180
Db 121 TFLIQLEBRKGVQSSGIMLTFWLVAVCALAIIRSKIMTALKEDAOYDLFDDITFYVYS 180
Qy 181 LLLIQVLVLSGSDSPFLFSETIHDNPNCPRESSASFLSRTFWMTGLIVRXYRQPLESSD 240
Db 181 LLLIQVLVLSGSDSPFLFSETIHDNPNCPRESSASFLSRTFWMTGLIVRXYRQPLESSD 240
Qy 241 LMSLNKEDTSQVVPVLYKMKKCAKTRKQPVVYVSSKDPAPKESKVDANEVVAL 300
Db 241 LMSLNKEDTSQVVPVLYKMKKCAKTRKQPVVYVSSKDPAPKESKVDANEVVAL 300
Qy 301 IVKSPQKEMWNSLFKVLTKTFGPYFLMSFFPKAIDIMFSGPOLKLLIKFVNDTKAPD 360
Db 301 IVKSPQKEMWNSLFKVLTKTFGPYFLMSFFPKAIDIMFSGPOLKLLIKFVNDTKAPD 360
Qy 361 WQGYFYTVLFLVTKACLOTVLVHOYFHHICFVSGMRKIKTAVIGAVYRKALVITNSARKSSTV 420
Db 361 WQGYFYTVLFLVTKACLOTVLVHOYFHHICFVSGMRKIKTAVIGAVYRKALVITNSARKSSTV 420

```

```

QY 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVIALYILMLNLSGVSILAGVAVMLAMPVN 480
DB 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVIALYILMLNLSGVSILAGVAVMLAMPVN 480
QY 481 AVMAKTKTYQVAMHMSKDNRIKLMNEILNGIKVLKLYAMELAFKDKVLAIROBELKVLK 540
DB 481 AVMAKTKTYQVAMHMSKDNRIKLMNEILNGIKVLKLYAMELAFKDKVLAIROBELKVLK 540
QY 541 KSAVLASVGTFTWCTPPLVALCTPAVYVITDENNILDQAQFVSLAFNLRPLNLP 600
DB 541 KSAVLASVGTFTWCTPPLVALCTPAVYVITDENNILDQAQFVSLAFNLRPLNLP 600
QY 601 MVISIVQASVSLKRLRIFLSHEELPDSIRRRPVKGGGNSITVRNATFTWARSDDPT 660
DB 601 MVISIVQASVSLKRLRIFLSHEELPDSIRRRPVKGGGNSITVRNATFTWARSDDPT 660
QY 661 LMGITFSPREGALVAVVGVCGKSSLSALLAEMDKVGHVAKGSVAVVPOQAWIOND 720
DB 661 LMGITFSPREGALVAVVGVCGKSSLSALLAEMDKVGHVAKGSVAVVPOQAWIOND 720
QY 721 SLRENILFGCOLLEPPYRSVIOACALLPDLILPSGDRTEIGEGKVLSSGQKQVSLAR 780
DB 721 SLRENILFGCOLLEPPYRSVIOACALLPDLILPSGDRTEIGEGKVLSSGQKQVSLAR 780
QY 781 AVVSNADIVLPDDLSAVDAVGHIFENVIGPKMKLKNKRIIVTHSMSTLPOVDYIV 840
DB 781 AVVSNADIVLPDDLSAVDAVGHIFENVIGPKMKLKNKRIIVTHSMSTLPOVDYIV 840
QY 841 MSGGKISEMGSYOELLARDGAFAPFLTYASTEOBDAEENGVTGSGPGKEAQOMENG 900
DB 841 MSGGKISEMGSYOELLARDGAFAPFLTYASTEOBDAEENGVTGSGPGKEAQOMENG 900
QY 901 LVTDSAGKOLOROUSSSSSYSGDISRHNSSTAELQKAEKKEFTWKLMEADKAQTGVKL 960
DB 901 LVTDSAGKOLOROUSSSSSYSGDISRHNSSTAELQKAEKKEFTWKLMEADKAQTGVKL 960
QY 961 SVVWDYKAKIQLFISFLIFLFCNHYVSALASNTWLSMTDDPIVNGTOETKRLSVYG 1020
DB 961 SVVWDYKAKIQLFISFLIFLFCNHYVSALASNTWLSMTDDPIVNGTOETKRLSVYG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLHVDLHSIRSPMSFEPTPSGULVNRFSKEL 1080
DB 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLHVDLHSIRSPMSFEPTPSGULVNRFSKEL 1080
QY 1081 DTVDMSIPEVYKMFMSGLFNVIACIYILLATPIAIIIPPLGLIYFVORFYVASSROL 1140
DB 1081 DTVDMSIPEVYKMFMSGLFNVIACIYILLATPIAIIIPPLGLIYFVORFYVASSROL 1140
QY 1141 KRLESVRSPPVSHFNFTLLGVSYIRAFEEBERFIHOSDKVDNOKRAYPSIYANWLA 1200
DB 1141 KRLESVRSPPVSHFNFTLLGVSYIRAFEEBERFIHOSDKVDNOKRAYPSIYANWLA 1200
QY 1201 VRLECVCNCIVLFAALFAVISRHSLASGLVGLSVSYSLQVTTYLNMVLRSSSEMETIVA 1260
DB 1201 VRLECVCNCIVLFAALFAVISRHSLASGLVGLSVSYSLQVTTYLNMVLRSSSEMETIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVFNRYCLRYREDLDFVLRHINVTNGG 1320
DB 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVFNRYCLRYREDLDFVLRHINVTNGG 1320
QY 1321 EKVGIVRTAGAGKSSLLGLFRINESAGEIITIDGINIAKIGLHDLFKTIIIPQDVL 1380
DB 1321 EKVGIVRTAGAGKSSLLGLFRINESAGEIITIDGINIAKIGLHDLFKTIIIPQDVL 1380
QY 1381 SGLSRMNLDPFSQYSDDEEVTSLSLAKDFVSALPKLDHBCAEGENISVGRQVL 1440
DB 1381 SGLSRMNLDPFSQYSDDEEVTSLSLAKDFVSALPKLDHBCAEGENISVGRQVL 1440
QY 1441 ARALLRKTKLIVDEATAVADLETDLIQSTIRFOFEDCTVLTAAHLNATIMDYTRIVL 1500
DB 1441 ARALLRKTKLIVDEATAVADLETDLIQSTIRFOFEDCTVLTAAHLNATIMDYTRIVL 1500

```

```

QY 1501 DKGEIOEYGAESPDLLOQRGLFYSMAXDAGLV 1531
DB 1501 DKGEIOEYGAESPDLLOQRGLFYSMAXDAGLV 1531

RESULT 9
AA78873
ID AA78873 standard; protein, 1531 AA.
XX
AC AA78873;
XX
DT 19-MAY-2000 (first entry)
XX
DE Multidrug resistance protein (MRP) natural variant amino acid sequence.
XX
KW Multidrug resistance protein; MRP; human; anthracycline; Vinca alkaloid;
XX
KW epipodophyllotoxin; cancer; leukaemia.
XX
OS Homo sapiens.
XX
EN US6025473-A.
XX
PD 15-FEB-2000.
XX
PE 05-JUN-1995; 95US-00461384.
XX
PR 27-OCT-1992; 92US-00966923.
XX
PR 08-MAR-1993; 93US-00029340.
XX
PR 26-OCT-1993; 93US-00141893.
XX
PR 20-MAR-1995; 95US-00407207.
XX
PA (TOOH ) UNIV QUEBENS KINGSTON.
XX
PI Cole SPC, Deeley RG;
XX
DR WPI: 2000-181838/16.
XX
DR N-PSDB; AA290193.
XX
PT Isolated protein conferring multidrug resistance, to at least two drugs
XX
PT selected from anthracyclines, epipodophyllotoxins and Vinca alkaloids, on
XX
PT a drug sensitive mammalian cell.
XX
PS Claim 10; Col 79-88; 78pp; English.
XX
CC This sequence represents a human multidrug resistance protein (MRP)
CC natural variant amino acid sequence. The human MRP confers multidrug
CC resistance, including resistance to at least two drugs selected from
CC anthracyclines, epipodophyllotoxins and Vinca alkaloids, on a drug
CC sensitive mammalian cell, when the protein is expressed in the cell. The
CC multidrug resistance is not substantially reversed by chemosensitizers
CC which reverse P-glycoprotein-mediated multidrug resistance. The MRP
CC protein sequence can be used to generate antibodies against MRP. The MRP
CC protein and nucleotide sequences can be used in compositions which are
CC used to treat patients with tumors displaying multidrug resistance. The
CC compositions and methods of the invention can be used particularly to
CC treat breast cancer, leukaemias, fibrosarcomas, cervical cancer, gliomas,
CC thymomas, neuroblastomas, and lung cancer. Antibodies directed against
CC MRP can be used to inhibit the multidrug resistance of a multidrug
CC resistant cell
XX
SQ Sequence 1531 AA:
SQ
Query Match 80.7%; Score 7860; DB 3; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGCSADGSPDLMDMNTVNTSNPDKCKONTLVVWPCPYLMACPEFYLSRH 60
DB 1 MALRGCSADGSPDLMDMNTVNTSNPDKCKONTLVVWPCPYLMACPEFYLSRH 60
QY 61 DRGYIOWTPLNKKTKTALGLFLMTVCWADLFSFWMERSGFIAPLAVPLVPTLGIITTLA 120
DB 61 DRGYIOWTPLNKKTKTALGLFLMTVCWADLFSFWMERSGFIAPLAVPLVPTLGIITTLA 120

```

QY 121 TFLIOLERRRGVOSSGIMLTFFMLVALYCALAIILRSKIMTALKEDAQVDLFEDITFFYYFS 180
Db 121 TFLIOLERRRGVOSSGIMLTFFMLVALYCALAIILRSKIMTALKEDAQVDLFEDITFFYYFS 180
QY 181 LLLIQLVLSCSDSDSPFSETIHDNCPRESSASFLSITTFWITGLIVRGYRQPLGSD 240
Db 181 LLLIQLVLSCSDSDSPFSETIHDNCPRESSASFLSITTFWITGLIVRGYRQPLGSD 240
QY 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKOPVKVYVSSKDPAPKRESSKVDABEVEAL 300
Db 241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKOPVKVYVSSKDPAPKRESSKVDABEVEAL 300
QY 301 IVKSPKEMNPSLFKVLYKTGPYFLMSFFPKAIHDLMMFSGPOLIKLLIKFVNDTRAPD 360
Db 301 IVKSPKEMNPSLFKVLYKTGPYFLMSFFPKAIHDLMMFSGPOLIKLLIKFVNDTRAPD 360
QY 361 WQGFYTVVLVFTVCLQTLVHOYFHICFVSGMIRKTAIVIGAVYRKALVITNSARKSTV 420
Db 361 WQGFYTVVLVFTVCLQTLVHOYFHICFVSGMIRKTAIVIGAVYRKALVITNSARKSTV 420
QY 421 GEIYNLMSVDAQRFMDLATYINMIWSAPLOVILALYLLMLNIGPSVLGAVAVMLMVPVN 480
Db 421 GEIYNLMSVDAQRFMDLATYINMIWSAPLOVILALYLLMLNIGPSVLGAVAVMLMVPVN 480
QY 481 AVMAKTKTYOVAMHKSXONRIKLMNEILNGIKVLKLYAMELAFKDYVLAIROBELKVLK 540
Db 481 AVMAKTKTYOVAMHKSXONRIKLMNEILNGIKVLKLYAMELAFKDYVLAIROBELKVLK 540
QY 541 KSAIILSAVGTFTWCTPFLVALCTPAVYVITDENNILLDAQTPASIALFNLIRPPNLTLP 600
Db 541 KSAIILSAVGTFTWCTPFLVALCTPAVYVITDENNILLDAQTPASIALFNLIRPPNLTLP 600
QY 601 MVISIYOASVSLKRLRIFLSHEBELPDSIRBRPKOOGGNSITVRNATFTMARSDPPT 660
Db 601 MVISIYOASVSLKRLRIFLSHEBELPDSIRBRPKOOGGNSITVRNATFTMARSDPPT 660
QY 661 LNGITFSIPEGALVAVVQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAMIOND 720
Db 661 LNGITFSIPEGALVAVVQVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAMIOND 720
QY 721 SLRENILFGCOLBEYRHSVIOACALLPDLBILPSGDBTEIGEKVNISSGQOKRVSLAR 780
Db 721 SLRENILFGCOLBEYRHSVIOACALLPDLBILPSGDBTEIGEKVNISSGQOKRVSLAR 780
QY 781 AVYENADLYLPDDPLSAVDAHVGKHIIFENVIGPKQMLKNKTRILVTHSMSTLPQVDVITV 840
Db 781 AVYENADLYLPDDPLSAVDAHVGKHIIFENVIGPKQMLKNKTRILVTHSMSTLPQVDVITV 840
QY 841 MSGKISEMGSYOELLARDGAFAEFLRTYASTEOBDAEENGVTGVSQPGKEAKOMENG 900
Db 841 MSGKISEMGSYOELLARDGAFAEFLRTYASTEOBDAEENGVTGVSQPGKEAKOMENG 900
QY 901 LVTTDSAGKOLOROLSSSSSYSGDISRRHNSSTBELOKAEAKKEETKMLEADKAOTGOYKL 960
Db 901 LVTTDSAGKOLOROLSSSSSYSGDISRRHNSSTBELOKAEAKKEETKMLEADKAOTGOYKL 960
QY 961 SVYMDYMAIGLFTSFLSIFLPMCHVASALSNWMLSLMTDDPIYNGOETHKVPLSYVG 1020
Db 961 SVYMDYMAIGLFTSFLSIFLPMCHVASALSNWMLSLMTDDPIYNGOETHKVPLSYVG 1020
QY 1021 ALGISOGLAVFGYSMAVASIGGILASRCLHVDLHSLILRSPMSFFERTSPGMLVNFSEKL 1080
Db 1021 ALGISOGLAVFGYSMAVASIGGILASRCLHVDLHSLILRSPMSFFERTSPGMLVNFSEKL 1080
QY 1081 DTVDMSIMEVIKMFNGSLFNVIAGCIVILLATPIAIIIPGLIYFFVQRPYVASRQOL 1140
Db 1081 DTVDMSIMEVIKMFNGSLFNVIAGCIVILLATPIAIIIPGLIYFFVQRPYVASRQOL 1140
QY 1141 KRLSVSRSPVYSHNFTLLGVSVIRAFEEORFTHQDIAKVDENOKAYYSBIVANRWLA 1200
Db 1141 KRLSVSRSPVYSHNFTLLGVSVIRAFEEORFTHQDIAKVDENOKAYYSBIVANRWLA 1200

QY 1201 VRLFCVNCIVLPAALFAVISRHSLSAGLVGSYSLSQVTTYINMLVRMSSEMETNIVA 1260
Db 1201 VRLFCVNCIVLPAALFAVISRHSLSAGLVGSYSLSQVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSENEKAPWQIOETAPSSWPQYGRVFRNYCLRYEDDLFPVLRHINVTINGG 1320
Db 1261 VERLKEYSENEKAPWQIOETAPSSWPQYGRVFRNYCLRYEDDLFPVLRHINVTINGG 1320
QY 1321 EKVYIGRTAGKSSLLTGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPVL 1380
Db 1321 EKVYIGRTAGKSSLLTGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPVL 1380
QY 1381 SCSLRNMLDPPSGYSDEEVTWTSLELAHLKDFVSALPKDLHCEAGGENSEVGOLOVCL 1440
Db 1381 SCSLRNMLDPPSGYSDEEVTWTSLELAHLKDFVSALPKDLHCEAGGENSEVGOLOVCL 1440
QY 1441 ABALARKTKILVDEAFAVADLETDDLIQSTIRTOFEDCTVLTFAHLANTIMDYTRIVL 1500
Db 1441 ABALARKTKILVDEAFAVADLETDDLIQSTIRTOFEDCTVLTFAHLANTIMDYTRIVL 1500
QY 1501 DKGEIOEYGAPSDILQQRGLFYSMAKDAGLV 1531
Db 1501 DKGEIOEYGAPSDILQQRGLFYSMAKDAGLV 1531

RESULT 10
ABG61810
ID ABG61810 standard; protein; 1531 AA.
XX
AC ABG61810;
XX
DT 15-AUG-2002 (first entry)
XX
DE Prostate cancer-associated protein #11.
XX
KM Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
XX
OS Mammalia.
XX
PN MO200230268-A2.
XX
PD 18-APR-2002.
XX
PE 12-OCT-2001; 2001WO-US032045.
XX
PR 13-OCT-2000; 2000US-00687576.
PR 08-DEC-2000; 2000US-00733388.
PR 08-DEC-2000; 2000US-00733742.
PR 24-JAN-2001; 2001US-0263957P.
PR 16-MAR-2001; 2001US-0276791P.
PR 06-APR-2001; 2001US-0276888P.
PR 06-APR-2001; 2001US-0281922P.
PR 24-APR-2001; 2001US-0286214P.
PR 30-APR-2001; 2001US-00847046.
PR 04-MAY-2001; 2001US-0288589P.
PA (BOSB-) BOS BIOTECHNOLOGY INC.
XX
PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX
DR WP1; 2002-471335/50.
DR N-PSDB; ABK92125.
XX
PT Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.
XX
PS Claim 27; Page 309; 436pp; English.
XX
CC The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with prostate cancer-

CC associated polynucleotides (designated PC genes) that selectively
CC hybridise to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins
XX
SQ Sequence 1531 AA:
Query Match 80.7%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSADGSDMLDMNVTWNTSNPDTKCPONTLVVWPCFYLMACPFYFLYLSRH 60
DB 1 MALRGFCSADGSDMLDMNVTWNTSNPDTKCPONTLVVWPCFYLMACPFYFLYLSRH 60
QY 61 DRGYIOMTPIKTKTALGFLIMIVCMADLFYSFWEBSRGIFLAVFLVSPFLGITTLA 120
DB 61 DRGYIOMTPIKTKTALGFLIMIVCMADLFYSFWEBSRGIFLAVFLVSPFLGITTLA 120
QY 121 TFLIQLERRRGVSGSGLMFLWALVALCALILRSKIMTALKEADAQVDFPDITFYVFS 180
DB 121 TFLIQLERRRGVSGSGLMFLWALVALCALILRSKIMTALKEADAQVDFPDITFYVFS 180
QY 181 LLLIQLVLSGCSDBSPFSETIHDNPNCPBSSASFLSRITRWITGILVGRYROPLEGS 240
DB 181 LLLIQLVLSGCSDBSPFSETIHDNPNCPBSSASFLSRITRWITGILVGRYROPLEGS 240
QY 181 LLLIQLVLSGCSDBSPFSETIHDNPNCPBSSASFLSRITRWITGILVGRYROPLEGS 240
DB 181 LLLIQLVLSGCSDBSPFSETIHDNPNCPBSSASFLSRITRWITGILVGRYROPLEGS 240
QY 241 LMSLNKEDTSEQVVPVVKWKKKCACTRKOPVAVVSSKOPAPCKSSKVDAAEEVAL 300
DB 241 LMSLNKEDTSEQVVPVVKWKKKCACTRKOPVAVVSSKOPAPCKSSKVDAAEEVAL 300
QY 301 IVKSPQKEMNPSELFVLYKTGPFYFLMSFFPKAIDHLMFSGPOLILKLIFVNDTRAP 360
DB 301 IVKSPQKEMNPSELFVLYKTGPFYFLMSFFPKAIDHLMFSGPOLILKLIFVNDTRAP 360
QY 361 WQGYFTYVLLFTACLOTVLHQYFHIQFVSGMRKTAIVGAYRKALVITNSARKSTV 420
DB 361 WQGYFTYVLLFTACLOTVLHQYFHIQFVSGMRKTAIVGAYRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDAORFMDLATYINMIWSAPLOYTLALYLLMLGSPVLAGVAVVWLMYPN 480
DB 421 GEIVNLMSVDAORFMDLATYINMIWSAPLOYTLALYLLMLGSPVLAGVAVVWLMYPN 480
QY 481 AVAMAKTKTYOVAMKSKNRKIKLMEILINGIKVLYLAMELAFKDKVLAIROBELKVLK 540
DB 481 AVAMAKTKTYOVAMKSKNRKIKLMEILINGIKVLYLAMELAFKDKVLAIROBELKVLK 540
QY 541 KSAVLAVGFTFWCTPFLVALCTFAVYVITDENNNIIDAOFAVSLFNLIRPPLILP 600
DB 541 KSAVLAVGFTFWCTPFLVALCTFAVYVITDENNNIIDAOFAVSLFNLIRPPLILP 600
QY 601 MVISSIVQASVSLKRLIFLSHELEPDSIERRPVKOGGGSNSTVNNAPFTARSPT 660
DB 601 MVISSIVQASVSLKRLIFLSHELEPDSIERRPVKOGGGSNSTVNNAPFTARSPT 660
QY 661 LNCITFESIPGALVAVVGVGCGKSSLSALLAMEKVEGHVALKGSVAVVPOQAWIOND 720
DB 661 LNCITFESIPGALVAVVGVGCGKSSLSALLAMEKVEGHVALKGSVAVVPOQAWIOND 720
QY 721 SLSENILFGQLEBPYRSVIOACALLPDLEILBPSGRTIIGEGVNLSSGQKORVSLAR 780
DB 721 SLSENILFGQLEBPYRSVIOACALLPDLEILBPSGRTIIGEGVNLSSGQKORVSLAR 780
QY 781 AYSNADIYLFDDPLSAVDAHVGKHI FENVYIGPKMLKNKTRILVTHSMYSLPQVDYIIV 840
DB 781 AYSNADIYLFDDPLSAVDAHVGKHI FENVYIGPKMLKNKTRILVTHSMYSLPQVDYIIV 840

DB 781 AYSNADIYLFDDPLSAVDAHVGKHI FENVYIGPKMLKNKTRILVTHSMYSLPQVDYIIV 840
QY 841 MSGGKISEMSYQELLARDGAFAEFLTITASTBOBDAENGVTGSGPGKEAKOMENGM 900
DB 841 MSGGKISEMSYQELLARDGAFAEFLTITASTBOBDAENGVTGSGPGKEAKOMENGM 900
QY 901 LMTDSACKOLQOROLSSSSSYSGDISRRHNSHTAELOKAEKKEETWKLMEDAKOTGVKL 960
DB 901 LMTDSACKOLQOROLSSSSSYSGDISRRHNSHTAELOKAEKKEETWKLMEDAKOTGVKL 960
QY 961 SYVWDYMKALGLFISFLIFLFCMNHVSALASNYMISLMTDDPIVNGTOEHTKRLSVYG 1020
DB 961 SYVWDYMKALGLFISFLIFLFCMNHVSALASNYMISLMTDDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISQGIAYFGYSMAVSGIILASRCLAVDLHSILRSFMSFPERTPSGNLVRFSEKL 1080
DB 1021 ALGISQGIAYFGYSMAVSGIILASRCLAVDLHSILRSFMSFPERTPSGNLVRFSEKL 1080
QY 1081 DTVDMSIPEVYIKPMFMSLFNVIGACIVILATPIAIIIPPLGLIFEFVORFYVASSROL 1140
DB 1081 DTVDMSIPEVYIKPMFMSLFNVIGACIVILATPIAIIIPPLGLIFEFVORFYVASSROL 1140
QY 1141 KRLESYSRSPVYSHFNFTLLGVSIVIRAFEEQERFIHQSDLKVDENQATYPSIVANWLA 1200
DB 1141 KRLESYSRSPVYSHFNFTLLGVSIVIRAFEEQERFIHQSDLKVDENQATYPSIVANWLA 1200
QY 1201 VRLBCVNCIULPALPAVVISRHSLSAGLVGSVSYSLQVTTYLNLVRRSSEMETNIVA 1260
DB 1201 VRLBCVNCIULPALPAVVISRHSLSAGLVGSVSYSLQVTTYLNLVRRSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMIOETAPSSWPQVGRVFRNYCIRYREDLDFVLRIHNTVING 1320
DB 1261 VERLKEYSETEKAPMIOETAPSSWPQVGRVFRNYCIRYREDLDFVLRIHNTVING 1320
QY 1321 EKYGIWRTGAGKSSLTGLFGRINESABGEIIDGIMIAKIGLHDLRFKTTIIPQDPVLF 1380
DB 1321 EKYGIWRTGAGKSSLTGLFGRINESABGEIIDGIMIAKIGLHDLRFKTTIIPQDPVLF 1380
QY 1381 SGLSRNMLDPSQYSDDEVWTSLELAHLKDFVSLPDKLHCECAGEGNSLVGQRLVCL 1440
DB 1381 SGLSRNMLDPSQYSDDEVWTSLELAHLKDFVSLPDKLHCECAGEGNSLVGQRLVCL 1440
QY 1441 ARALLRKTILVLDDEATAVNDLSTDLIOSTIRTOFEDCTVLITARLNTIMDYTRIVL 1500
DB 1441 ARALLRKTILVLDDEATAVNDLSTDLIOSTIRTOFEDCTVLITARLNTIMDYTRIVL 1500
QY 1501 DKGEIOEYGA PSDLLOORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGA PSDLLOORGLFYSMAKDAGLV 1531
RESULT 11
ID AEM35012
ID AEM35012 standard; protein; 1531 AA.
XX AEM35012;
AC XX
AC XX
DT 08-OCT-2003 (first entry)
XX
DE Cancer based on CYP3A5 related protein SEQ ID NO:678.
XX
XX Cancer; CYP3A5; irinotecan; pharmaceutical; malignant glioma;
KM cytochrome p450; subfamily IIIA; nifedipine oxidase; polypeptide 5;
XX cytostatic.
OS Unidentified.
XX
PN W02003013534-A2.
XX
XX 20-FEB-2003.
PD
XX
PF 23-JUL-2002; 2002MO-BF008219.
XX

PR 23-JUL-2001; 2001EP-00117608.
PR 24-MAY-2002; 2002EP-00011710.
XX
PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI Heinrich G, Kern R;
XX WPI; 2003-268144/26.
DR
XX
PT New use of irinotecan for preparation of compositions for treating cancer
PT in subject having genome with variant allele comprising cytochrome p450,
PT subfamily IIIA, polypeptide 5 polynucleotide, termed CYP3A5.
XX
XX Disclosure; SEQ ID NO 678; 86pp; English.
XX
CC The present invention describes the use of irinotecan (I) or its
CC derivative for the preparation of a pharmaceutical composition for
CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic
CC cancer, or malignant glioma in a subject having a genome with a variant
CC allele, which comprises a cytochrome p450, subfamily IIIA (nifedipine
CC oxidase), polypeptide 5 (CYP3A5) polynucleotide (II). (I) and (II) have
CC cytostatic activity. The therapeutic applications of (I) is improved,
CC since it is possible to individually treat a subject with an appropriate
CC dosage and/or an appropriate derivative of (I). Therefore, undesirable,
CC harmful or toxic effects are efficiently avoided. Unnecessary and
CC potentially harmful treatment of those subjects who do not respond to the
CC treatment with substances (nonresponders), as well as the development of
CC drug resistances due to suboptimal drug dosing can be avoided. ACF62200
CC to ACF62751 and ABM34912 to ABM35013 represent sequences used in the
CC exemplification of the present invention
CC
XX
SQ Sequence 1531 AA;
Query Match 80.7%; Score 7860; DB 6; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCGADGSDPLMDNMVNTWNTSNPDFTKCFONTLVWVPCFYLMACPPFFLYLSRH 60
DB 1 MALRGFCGADGSDPLMDNMVNTWNTSNPDFTKCFONTLVWVPCFYLMACPPFFLYLSRH 60
QY 61 DRGTYQMPLNKTALGFLIMIVCMADLFYSFWRSGITFLAPFLVSPFLTGTTLLA 120
DB 61 DRGTYQMPLNKTALGFLIMIVCMADLFYSFWRSGITFLAPFLVSPFLTGTTLLA 120
QY 121 TFLQLERKRGVSSGIMLTFMLVALVCLALILRSKIMTALKEQOVDFRDITFYVFS 180
DB 121 TFLQLERKRGVSSGIMLTFMLVALVCLALILRSKIMTALKEQOVDFRDITFYVFS 180
QY 181 LLLIQLVLSCFSDRSPLFSETIHDNPPCESSASFLSRITFMWITGLIVRGYRPLEGSD 240
DB 181 LLLIQLVLSCFSDRSPLFSETIHDNPPCESSASFLSRITFMWITGLIVRGYRPLEGSD 240
QY 241 LWSLNKRETSBOVYVLYVMKKKECAKTRKOPVKVYVSSKDPADOKESSKYDANEVVAL 300
DB 241 LWSLNKRETSBOVYVLYVMKKKECAKTRKOPVKVYVSSKDPADOKESSKYDANEVVAL 300
QY 301 IVKSPKEMNPSLFLVLYKTFGPFLYMSFFPKAIDHLMFSGPOLIKLLIKFVNDTKAPD 360
DB 301 IVKSPKEMNPSLFLVLYKTFGPFLYMSFFPKAIDHLMFSGPOLIKLLIKFVNDTKAPD 360
QY 361 WQGYFFYVLVFTVLAQTLVLHOFYFHCIFVSGMRITKAVIGAIVRKALVITNSARKSSTV 420
DB 361 WQGYFFYVLVFTVLAQTLVLHOFYFHCIFVSGMRITKAVIGAIVRKALVITNSARKSSTV 420
QY 421 GEIYNLMSVDAORFMDLATYINMTSAPLOVTLAYLLMLNIGSEVLAGVAVMLVMEVN 480
DB 421 GEIYNLMSVDAORFMDLATYINMTSAPLOVTLAYLLMLNIGSEVLAGVAVMLVMEVN 480
QY 481 AVMAKTKTYOVAAHMSKDNRIKLMNEITLNGIKVLKLYAMELAFDKVLAIRQELKYLK 540
DB 481 AVMAKTKTYOVAAHMSKDNRIKLMNEITLNGIKVLKLYAMELAFDKVLAIRQELKYLK 540

QY 541 KSAIYLSAVGFTTWCTPFLVALCTFPAVYVITDENNNIIDAOCTAFSLAFNLRPEPLNLP 600
DB 541 KSAIYLSAVGFTTWCTPFLVALCTFPAVYVITDENNNIIDAOCTAFSLAFNLRPEPLNLP 600
QY 601 MVISSIVQASVSLRLRIPLSHEELAPDSTERPVKDGGSNTSTVNAATTFMARSDEPT 660
DB 601 MVISSIVQASVSLRLRIPLSHEELAPDSTERPVKDGGSNTSTVNAATTFMARSDEPT 660
QY 661 LMGITFSIPGALVAVVQVCGKSSLLSALLAEMDKVEGHVALKGSVAVVPOQAMIOND 720
DB 661 LMGITFSIPGALVAVVQVCGKSSLLSALLAEMDKVEGHVALKGSVAVVPOQAMIOND 720
QY 721 SLRENILFCCQLPEPPTYSVYQAALLPDLEILPSGRTGEGKGNLSSGQKQVSLAR 780
DB 721 SLRENILFCCQLPEPPTYSVYQAALLPDLEILPSGRTGEGKGNLSSGQKQVSLAR 780
QY 781 AVYSNADITYLPDDPLSAVDAGKHIFENNYGPRKMLKNKRIIVTHSMSTLPQVDVYIV 840
DB 781 AVYSNADITYLPDDPLSAVDAGKHIFENNYGPRKMLKNKRIIVTHSMSTLPQVDVYIV 840
QY 841 MSGGKISMSGYQELRLARDGAFAEFLRTYASTBOEODAEENGVTGSGPGKEAKOMENG 900
DB 841 MSGGKISMSGYQELRLARDGAFAEFLRTYASTBOEODAEENGVTGSGPGKEAKOMENG 900
QY 901 LVYTSAGKQLOROLSSSSSYSGDISRHNNSTAELOKAKKEETWKLMEADKAOQGVYL 960
DB 901 LVYTSAGKQLOROLSSSSSYSGDISRHNNSTAELOKAKKEETWKLMEADKAOQGVYL 960
QY 961 SVYDYMKAIGLFLSPISIFLFMCNHYSAASNYWLSMTDPIVNGTOETHKRLSYG 1020
DB 961 SVYDYMKAIGLFLSPISIFLFMCNHYSAASNYWLSMTDPIVNGTOETHKRLSYG 1020
QY 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLHLSILRSPSPFERTPSGNLVNRFSKEL 1080
DB 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLHLSILRSPSPFERTPSGNLVNRFSKEL 1080
QY 1081 DTVDSMILPEVTKMWSLFNVYIGACIYILATPLAIIIPPLGLIYFFVQFYVASRQL 1140
DB 1081 DTVDSMILPEVTKMWSLFNVYIGACIYILATPLAIIIPPLGLIYFFVQFYVASRQL 1140
QY 1141 KRLSVSRSPVYSHFNETLGVSVIYRAFEQERFIHOSDLKVDENOKAYYSIVANRMLA 1200
DB 1141 KRLSVSRSPVYSHFNETLGVSVIYRAFEQERFIHOSDLKVDENOKAYYSIVANRMLA 1200
QY 1201 VRLCEVNGCIYLFALPAVISRHSLSAGLVGLSVYSLOVTTYINMLVRSSEMETNIVA 1260
DB 1201 VRLCEVNGCIYLFALPAVISRHSLSAGLVGLSVYSLOVTTYINMLVRSSEMETNIVA 1260
QY 1261 VERLKEYSFTEKEAPWQIOETAPPSWPQVGRVEFRNYCLAYRDLDFVLHINVTINGG 1320
DB 1261 VERLKEYSFTEKEAPWQIOETAPPSWPQVGRVEFRNYCLAYRDLDFVLHINVTINGG 1320
QY 1321 EKVQIVRGTAGKSLTLGLPRINESAGEIIIDIGINAKIGLHDLRKTIIIPQDPFLF 1380
DB 1321 EKVQIVRGTAGKSLTLGLPRINESAGEIIIDIGINAKIGLHDLRKTIIIPQDPFLF 1380
QY 1381 SSGSLRMLNDPPFSQVSDSEVWTSLELAHLKDFVSAALPKLDHECAEGENLSVGOQOLVCL 1440
DB 1381 SSGSLRMLNDPPFSQVSDSEVWTSLELAHLKDFVSAALPKLDHECAEGENLSVGOQOLVCL 1440
QY 1441 ARALLRKTIIIVLEDAVTAVDLETDLDIQTIRFQEDCTVLTIAHRLNTIMDYTRIVL 1500
DB 1441 ARALLRKTIIIVLEDAVTAVDLETDLDIQTIRFQEDCTVLTIAHRLNTIMDYTRIVL 1500
QY 1501 DKGEIOEYGASDILQORGLFYSAKXAGLV 1531
DB 1501 DKGEIOEYGASDILQORGLFYSAKXAGLV 1531
RESULT 12
ADB20865
ID ADB20865 standard; protein; 1531 AA.
XX

Db 1381 SGRIRMLNDPESQYSDBEWTSLELAHLKDFVSALPDKLDEHCABGGENLSVQGRQLVCL 1440
QY 1441 ARALIRKTKILVLEATPAVDLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVVL 1500
Db 1441 ARALIRKTKILVLEATPAVDLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVVL 1500
QY 1501 DKGEIQEYGAPSDLLQGRGLFYSMKADGLV 1531
Db 1501 DKGEIQEYGAPSDLLQGRGLFYSMKADGLV 1531
RESULT 13
ADB87954
ID ADB87954 standard; protein; 1531 AA.
XX ADB87954;
AC ADB87954;
XX 04-DEC-2003 (first entry)
DT 04-DEC-2003 (first entry)
XX Human UGT1A1 protein sequence SEQ ID NO:678.
XX
XX irinotecan; cancer; UGT1A1; cytostatic; topoisomerase I inhibitor;
XX colorectal cancer; cervical cancer; gastric cancer; lung cancer;
XX ovarian cancer; pancreatic cancer; malignant glioma;
XX uridine diphosphate glycosyltransferase1 member A1.
XX
XX Homo sapiens.
XX
XX WO2003013536-A2.
XX
XX 20-FEB-2003.
XX
XX 23-JUL-2002; 2002WO-EP008217.
XX
XX 23-JUL-2001; 2001EP-00117608.
XX
XX 24-MAY-2002; 2002EP-00011710.
XX
XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
XX Heinrich G, Kerb R;
XX
XX WPI; 2003-289896/28.
XX
XX Use of irinotecan to treat cancer patient by determining if patient has
XX variant alleles of UGT1A1 gene, administering increased/decreased amounts
XX of irinotecan based on increased/decreased levels of UGT1A1 gene product.
XX
XX Disclosure; SEQ ID NO 678; 107pp; English.
XX
XX The invention relates to the novel use of irinotecan to treat a patient
XX suffering from cancer. This involves determining if the patient has one
XX or more variant alleles of the UGT1A1 gene, and if the patient has one
XX or more of such variant alleles, irinotecan is administered in an increased
XX or decreased amount in comparison to the amount that is administered
XX without regard to the patient's alleles in the UGT1A1 gene. The invention
XX has cytostatic activity. A composition of the invention acts as a
XX topoisomerase I inhibitor. The method is useful for treating a patient.
XX an animal e.g. mouse or a human, preferably African or Asian, suffering
XX from cancer such as colorectal, cervical, gastric cancer, lung, ovarian,
XX pancreatic cancer or malignant glioma. The present sequence is udes in
XX the exemplification of the invention.
XX
XX Sequence 1531 AA:
SQ
Query Match 80.7%; Score 7860; DB 7; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 61 DRGYQMTPLNKTKTALGFLMLIVCMADLFYSFWERSRGIFLAPVFLVSPTLIGITTTLLA 120
QY 121 TFLIQERRRGVSSGIMLTFWLVAVALCALAIRSKIMTALKEDAQVDLPEDITFYVYFS 180
Db 121 TFLIQERRRGVSSGIMLTFWLVAVALCALAIRSKIMTALKEDAQVDLPEDITFYVYFS 180
QY 181 LLLIQVLVSCFSDBSPLESETIHDNPPCPSSASFLSRITFWMITGLIVRGYRQPLEGSD 240
Db 181 LLLIQVLVSCFSDBSPLESETIHDNPPCPSSASFLSRITFWMITGLIVRGYRQPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLYKWKKECATRKQPVVYVSSQDPAPPKSSSKYDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLYKWKKECATRKQPVVYVSSQDPAPPKSSSKYDANEVEAL 300
QY 301 IVKSPKEMNPSSLFKVLKYTFGPYELMSFPFKAHDLMMFSGPOILKLLIKFVNDTPRAPD 360
Db 301 IVKSPKEMNPSSLFKVLKYTFGPYELMSFPFKAHDLMMFSGPOILKLLIKFVNDTPRAPD 360
QY 361 WQGYFTVLLFVTAQTLQTLVHQYFHI CPVSGMKIKTAVIGAVYRKALVITNSARKSTV 420
Db 361 WQGYFTVLLFVTAQTLQTLVHQYFHI CPVSGMKIKTAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLMSSVDAQRFMDLATTYINMTWSAPLOYITALLYLMNLGFSYLAGAVVMTLMPVN 480
Db 421 GEIVNLMSSVDAQRFMDLATTYINMTWSAPLOYITALLYLMNLGFSYLAGAVVMTLMPVN 480
QY 481 AVNMAKTKTOVAVHMKSKDNRIKLMNEITNGIKYKLYAMELAFKDKVLAIROEELKYLK 540
Db 481 AVNMAKTKTOVAVHMKSKDNRIKLMNEITNGIKYKLYAMELAFKDKVLAIROEELKYLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTPAVYVITDENNIIDAQTAFLVALFNILRPPLNLT 600
Db 541 KSAVLSAVGTFTWCTPFLVALCTPAVYVITDENNIIDAQTAFLVALFNILRPPLNLT 600
QY 601 MVISIYQASVSLKRLKIFLSHELEPDSIERRKYVQGGGNTSTVNATFTWARSDEPT 660
Db 601 MVISIYQASVSLKRLKIFLSHELEPDSIERRKYVQGGGNTSTVNATFTWARSDEPT 660
QY 661 LMGITFSPREGALAAVVGQVCGKSSLLSALLAEMDKVEGHVALIKGSAVYVPOQAMIOND 720
Db 661 LMGITFSPREGALAAVVGQVCGKSSLLSALLAEMDKVEGHVALIKGSAVYVPOQAMIOND 720
QY 721 SLRENIIFGCGLEPPYRSVIOACALLPDLIELPSGRTIEGKGVNLSGGQKORVSLAR 780
Db 721 SLRENIIFGCGLEPPYRSVIOACALLPDLIELPSGRTIEGKGVNLSGGQKORVSLAR 780
QY 781 AVYSNADITYLFDDPLSAVDHAVGKIFBNVIGPKMLKNKTRIIVTHSMSTLPOVDVITV 840
Db 781 AVYSNADITYLFDDPLSAVDHAVGKIFBNVIGPKMLKNKTRIIVTHSMSTLPOVDVITV 840
QY 841 MSGKISPMGSYOELARDGAFAEFLRTYASTBOEADAENGVGVSGPGEAKOMENGM 900
Db 841 MSGKISPMGSYOELARDGAFAEFLRTYASTBOEADAENGVGVSGPGEAKOMENGM 900
QY 901 LVTSAGKQOROLSSSSSYSGDISRHNSYAELOKAEKKEETWKLMEADKAOTGYKL 960
Db 901 LVTSAGKQOROLSSSSSYSGDISRHNSYAELOKAEKKEETWKLMEADKAOTGYKL 960
QY 961 SVYWDYMKAIGLFISLISIFLPMCNHVSALASNYWLSMTDDPIVNGTOEHTKYRLSYG 1020
Db 961 SVYWDYMKAIGLFISLISIFLPMCNHVSALASNYWLSMTDDPIVNGTOEHTKYRLSYG 1020
QY 1021 ALGISQGIYAVRGYMAVSIIGIILASRCHVDLHLSIIRSPMSFERTSGULVNRFSKL 1080
Db 1021 ALGISQGIYAVRGYMAVSIIGIILASRCHVDLHLSIIRSPMSFERTSGULVNRFSKL 1080
QY 1081 DTVDSMIPEVIMKMGSLFNIVIGACIVILATPIAIIIPPLGILYFFVQGFYVASSROL 1140
Db 1081 DTVDSMIPEVIMKMGSLFNIVIGACIVILATPIAIIIPPLGILYFFVQGFYVASSROL 1140
QY 1141 KRLESVSRSPYSHNETLLGVSVIRAEQERPIHOSDLKVDENOKAYYSIVANRWLA 1200
Db 1141 KRLESVSRSPYSHNETLLGVSVIRAEQERPIHOSDLKVDENOKAYYSIVANRWLA 1200

Db 1441 KRLESVSRSPYSHFNELLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANWMLA 1200
QY 1201 VRLECVGNCIYLPALFAVISRHSLSAGLVSVYSYQVTTYLNTWLVMSSEMETNIVA 1260
Db 1201 VRLECVGNCIYLPALFAVISRHSLSAGLVSVYSYQVTTYLNTWLVMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPPSWMPQVGRVERFRNYCLVRREDLPVLNHNVTNGG 1320
Db 1261 VERLKEYSETEKEAPWQIOETAPPSWMPQVGRVERFRNYCLVRREDLPVLNHNVTNGG 1320
QY 1321 EKVGIVGTGAGKSSLTGLFRINBSAGEIILIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
Db 1321 EKVGIVGTGAGKSSLTGLFRINBSAGEIILIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGSRFMMULDPFSQSDSEWVTSLELAHKDFVSALPDLDEHCEAGGENLSVGROVLVCL 1440
Db 1381 SGSRFMMULDPFSQSDSEWVTSLELAHKDFVSALPDLDEHCEAGGENLSVGROVLVCL 1440
QY 1441 ARALLRKTKILVLEATNAVDLETTDDLIOSTIRTOFEDCTVLTJAHRLNTIMDYTRYVL 1500
Db 1441 ARALLRKTKILVLEATNAVDLETTDDLIOSTIRTOFEDCTVLTJAHRLNTIMDYTRYVL 1500
QY 1501 DKGEIQEYGAPSDLLQGRGLFYSMAXDAGLV 1531
Db 1501 DKGEIQEYGAPSDLLQGRGLFYSMAXDAGLV 1531

RESULT 14
ADB96937 standard; protein; 1531 AA.
ADB96937;
04-DEC-2003 (first entry)
Human MDR1 related protein sequence SEQ ID NO:678.
irinotecan; colorectal cancer; cervical cancer; gastric cancer;
lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
multidrug resistance 1; MDR1; cytosolic; human; Cyp3A5; MRP1; MDR1;
TOP1.
Homo sapiens.
MO2003013537-A2.
20-FEB-2003.
23-JUL-2002; 2002WO-EP008218.
23-JUL-2001; 2001EP-00117608.
24-MAY-2002; 2002EP-00011710.
(EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
Heinrich G, Kerb R;
WPI; 2003-268145/26.
New use of irinotecan for preparation of pharmaceutical compositions for
treating cancer in subject having genome with variant allele comprising
multidrug resistance 1 polynucleotide.
Disclosure; SEQ ID NO 678; 130pp; English.

CC glioma in a subject (preferably human, more preferably African or Asian)
CC or a mouse. The present sequence is used in the exemplification of the
CC invention.
XX Sequence 1531 AA;
SQ
Query Match 80.7%; Score 7860; DB 7; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSDADSDPLMDMNTWMTSNPDTKCFQNTLVWVPCFYLMACEPFYLYSRH 60
Db 1 MALRGFCSDADSDPLMDMNTWMTSNPDTKCFQNTLVWVPCFYLMACEPFYLYSRH 60
QY 61 DRGIVQMTPLNKTALGFLMIVCAVDLYSPFESRGRGFLAFVFLVSPLLGITTLLA 120
Db 61 DRGIVQMTPLNKTALGFLMIVCAVDLYSPFESRGRGFLAFVFLVSPLLGITTLLA 120
QY 121 TFLIQLERRKGVQSSGIMLFWLVALCALAIRSKIMTLKEDAQVDLFRDITFYVFS 180
Db 121 TFLIQLERRKGVQSSGIMLFWLVALCALAIRSKIMTLKEDAQVDLFRDITFYVFS 180
QY 181 LLLIQLVLSGFSRSPLESTIHDNPNCPSSASFLSRITFWWITGLIVRGYRQPLEGSD 240
Db 181 LLLIQLVLSGFSRSPLESTIHDNPNCPSSASFLSRITFWWITGLIVRGYRQPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVIVKWKKECAKTRQPVKVVYSSKDPAQKESKVDANEVVAL 300
Db 241 LMSLNKEDTSEQVVPVIVKWKKECAKTRQPVKVVYSSKDPAQKESKVDANEVVAL 300
QY 301 IVKSPQKEMNPSEKVLKTFGPFLMSFPFKAIHDMFSGPOLIKLVNDTRAPD 360
Db 301 IVKSPQKEMNPSEKVLKTFGPFLMSFPFKAIHDMFSGPOLIKLVNDTRAPD 360
QY 361 WQGYFTVLLFVTAQOTLVLAHQYFHICFSGMRKIKTAVIGAVRKALVITNSARKSSTV 420
Db 361 WQGYFTVLLFVTAQOTLVLAHQYFHICFSGMRKIKTAVIGAVRKALVITNSARKSSTV 420
QY 421 GEIVNLMVSVDAPRMDIATYINNIWSAPLOVILLAYLMLNLPBSVLGAVAVMLVAVN 480
Db 421 GEIVNLMVSVDAPRMDIATYINNIWSAPLOVILLAYLMLNLPBSVLGAVAVMLVAVN 480
QY 481 AVNAMKTKTYQVAMHMSKDNRIKLMNEILNGIVLKYAMELAFKDVLAIROBELVLK 540
Db 481 AVNAMKTKTYQVAMHMSKDNRIKLMNEILNGIVLKYAMELAFKDVLAIROBELVLK 540
QY 541 KSAVLNAGVTFWCTPFLVALCTPAVYVITIDENNIIDAQTAFLSLFNLRFPLNLP 600
Db 541 KSAVLNAGVTFWCTPFLVALCTPAVYVITIDENNIIDAQTAFLSLFNLRFPLNLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHELEBDSIERRPVKDGGGTNSITVNNATFTWARSDDPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHELEBDSIERRPVKDGGGTNSITVNNATFTWARSDDPT 660
QY 661 LNCITTSIPRGALVAVWGVGCGKSSILSLALLAEMKVESHVAIKSVAVVPOAWI QND 720
Db 661 LNCITTSIPRGALVAVWGVGCGKSSILSLALLAEMKVESHVAIKSVAVVPOAWI QND 720
QY 721 SLRENILFGCQLEEPYRSYIOACALLPDEILIPSGDRTIEGKGVNLSCGOKORVSLAR 780
Db 721 SLRENILFGCQLEEPYRSYIOACALLPDEILIPSGDRTIEGKGVNLSCGOKORVSLAR 780
QY 781 AVYSNADIVLFDPLSLAVDAVHVKHIFENVIGKGMKNKTRILVTHSMYSYLPQVDVIIV 840
Db 781 AVYSNADIVLFDPLSLAVDAVHVKHIFENVIGKGMKNKTRILVTHSMYSYLPQVDVIIV 840
QY 841 MSGGKISEMKSVOELLARDAFAEPLRTASTQEDDAEENGVTVSGPKKAKQWENEM 900
Db 841 MSGGKISEMKSVOELLARDAFAEPLRTASTQEDDAEENGVTVSGPKKAKQWENEM 900
QY 901 LVYDSAGKQDQRLDSSSSYSYSGDISRHNSSTABLOAAEAKKETWKLMEADKQGTQGVKL 960
Db 901 LVYDSAGKQDQRLDSSSSYSYSGDISRHNSSTABLOAAEAKKETWKLMEADKQGTQGVKL 960

QY 961 SVYWDYKAIGLFLISFLIFLFCMCHVSALASNYWLSMTDPIYNGTOEHTKVLRSYVG 1020
DB 961 SVYWDYKAIGLFLISFLIFLFCMCHVSALASNYWLSMTDPIYNGTOEHTKVLRSYVG 1020
QY 1021 ALGISGSIYAVGYMANSYIGSILASRCHVDLHLSILRSPMSFPERPTSSGULVNRFSKEL 1080
DB 1021 ALGISGSIYAVGYMANSYIGSILASRCHVDLHLSILRSPMSFPERPTSSGULVNRFSKEL 1080
QY 1081 DTYSMTPEVTKMGMGSLFNVIYACTIVLLATPIAIIIPPLGLYFPVQSPYVASSRQL 1140
DB 1081 DTYSMTPEVTKMGMGSLFNVIYACTIVLLATPIAIIIPPLGLYFPVQSPYVASSRQL 1140
QY 1141 KRLESVSRSPYSHFNETLLGVSVIRAFEOERFIHOSDLKVDENOKAYYSIVANRWLA 1200
DB 1141 KRLESVSRSPYSHFNETLLGVSVIRAFEOERFIHOSDLKVDENOKAYYSIVANRWLA 1200
QY 1201 VRLCEVGNCTVLPALPAVIRSHSISAGLVGLSVSYSLQVTTYLNMVLRMSSEMTNIVA 1260
DB 1201 VRLCEVGNCTVLPALPAVIRSHSISAGLVGLSVSYSLQVTTYLNMVLRMSSEMTNIVA 1260
QY 1261 VERLKESETEKEAPMOIQETAPSPSPQVGRVREFRNYCLRYBEDLDVLRHINVTINGG 1320
DB 1261 VERLKESETEKEAPMOIQETAPSPSPQVGRVREFRNYCLRYBEDLDVLRHINVTINGG 1320
QY 1321 EKVGIVGRTGAGKSLTLGLFRINESAAGEIIIDGINIAKIGLHDLRPKITIIIPDPVLF 1380
DB 1321 EKVGIVGRTGAGKSLTLGLFRINESAAGEIIIDGINIAKIGLHDLRPKITIIIPDPVLF 1380
QY 1381 SGSLRPMNLDPSPQSYDEEVMWTSLELAHLKDFVSALPDLDHECAGGNNLSVGQRQVLC 1440
DB 1381 SGSLRPMNLDPSPQSYDEEVMWTSLELAHLKDFVSALPDLDHECAGGNNLSVGQRQVLC 1440
QY 1441 ARALLRKTILVLEBATAVLDLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYRIVYL 1500
DB 1441 ARALLRKTILVLEBATAVLDLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYRIVYL 1500
QY 1501 DKGEIOEYGAPSDLLQGRGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGAPSDLLQGRGLFYSMAKDAGLV 1531

RESULT 15
ADB92128
ID ADB92128 standard; protein; 1531 AA.
XX
AC ADB92128;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human MDRI related protein sequence SEQ ID NO:678.
XX
KW irinotecan; colorectal cancer; cervical cancer; gastric cancer;
KW lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
KW multidrug resistance 1; MDRI; cytosolic; human; UGT1A1; MRP1; Top1.
XX
OS Homo sapiens.
XX
PN WO2003013535-A2.
XX
PD 20-FEB-2003.
XX
PF 23-JUL-2002; 2002WO-EP008220.
XX
PR 23-JUL-2001; 2001EP-00117608.
PR 24-MAY-2002; 2002EP-00011710.
XX
PA (EPID-1) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI Heinrich G, Kerb R;
XX
DR WPI; 2003-342400/32.
XX

PT New use of irinotecan for preparation of pharmaceutical compositions for
PT treating cancer in subject having genome with variant allele comprising
PT multidrug resistance 1 polynucleotide.
XX
XX
PS Disclosure; SEQ ID NO 678; 1049p; English.
CC
CC The invention relates to a novel use of irinotecan or its derivative for
CC the preparation of a pharmaceutical composition for treating colorectal,
CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant
CC glioma in a subject having a genome with a variant allele which comprises
CC a multidrug resistance 1 (MDR1) polynucleotide. A composition of the
CC invention has cytostatic activity. The present sequence is used in the
CC exemplification of the invention.
XX
XX
SQ Sequence 1531 AA;
Query Match 80.7%; Score 7860; DB 7; Length 1531;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSADGSDPLMDMNVNTMNTSNPDTKCPQNTVLVWPFCYLMACPPPEFLYLSRH 60
DB 1 MALRGFCSADGSDPLMDMNVNTMNTSNPDTKCPQNTVLVWPFCYLMACPPPEFLYLSRH 60
QY 61 DRGYIOMTPLNKTKTALGFLIMVCMADLFYSFWEBSRGIFLAVFLVSPFLIGITTLA 120
DB 61 DRGYIOMTPLNKTKTALGFLIMVCMADLFYSFWEBSRGIFLAVFLVSPFLIGITTLA 120
QY 121 TELIQLERRKGVSSGIMLTFMLVALVCALALNSKIMTALKEBAQVDLPFDIFPYVFS 180
DB 121 TELIQLERRKGVSSGIMLTFMLVALVCALALNSKIMTALKEBAQVDLPFDIFPYVFS 180
QY 181 LLLIQVLVSCFSDSPFLFSETIHDNPNCPRESSAFLSRIFFMWITGLIVRGYROPLESGD 240
DB 181 LLLIQVLVSCFSDSPFLFSETIHDNPNCPRESSAFLSRIFFMWITGLIVRGYROPLESGD 240
QY 241 LMSLNKEDTSEQVVPVLVKKMKCEAKTRKOPVRYVYSSKDPAPKSSSKDANEVEVAL 300
DB 241 LMSLNKEDTSEQVVPVLVKKMKCEAKTRKOPVRYVYSSKDPAPKSSSKDANEVEVAL 300
QY 301 IVKSPQKEMNPSLTKVLKTFGPFLMSFFPKAHDLMFSGPOLKLLIFVNDTKAPD 360
DB 301 IVKSPQKEMNPSLTKVLKTFGPFLMSFFPKAHDLMFSGPOLKLLIFVNDTKAPD 360
QY 361 WQGYFTVLVFTVACLOTLVHQYFHI CFVSGMBIKTAVGAVVRKALVITNSARKSSTV 420
DB 361 WQGYFTVLVFTVACLOTLVHQYFHI CFVSGMBIKTAVGAVVRKALVITNSARKSSTV 420
QY 421 GEIVNLSVDAQRMDLATYTNMTWSAPLOVITALLYLMLNLGSPVLGAVAVMVLMPVN 480
DB 421 GEIVNLSVDAQRMDLATYTNMTWSAPLOVITALLYLMLNLGSPVLGAVAVMVLMPVN 480
QY 481 AVMAKTKTYOVAMKSKDNRIKLMNETLNCIKYLKYAMELAFKDKYLAIROELKYLK 540
DB 481 AVMAKTKTYOVAMKSKDNRIKLMNETLNCIKYLKYAMELAFKDKYLAIROELKYLK 540
QY 541 KSAVLSAVGTFWCTPFLVALCTFAVYVTIDENNII DAOAFYSALFNILRPPLNLP 600
DB 541 KSAVLSAVGTFWCTPFLVALCTFAVYVTIDENNII DAOAFYSALFNILRPPLNLP 600
QY 601 MVISISIVASVSLRLRIFLSHEELPDSIRRPVKOGGINSITVRNATFTWASDPT 660
DB 601 MVISISIVASVSLRLRIFLSHEELPDSIRRPVKOGGINSITVRNATFTWASDPT 660
QY 661 LINGTFSIPBEGALVAVGQVCGKSSILSALLAEMDKVGHVALKGSAAVYPQQAQWIND 720
DB 661 LINGTFSIPBEGALVAVGQVCGKSSILSALLAEMDKVGHVALKGSAAVYPQQAQWIND 720
QY 721 SLRENILFGCOLBEPYRSVIOACALLPDLIELPSGDRTEIGEGVNLSGGOKQSVSLAR 780
DB 721 SLRENILFGCOLBEPYRSVIOACALLPDLIELPSGDRTEIGEGVNLSGGOKQSVSLAR 780
QY 781 AVYSNADITYLPDPLSAVDAAHVGRHIFENYIGPKGMLKNKTRILVTHSMSTYLPQVDVITV 840

Db 781 AAVSNADIIYLFDDPLSAVDAAHVKHII FENVIGPKGMLKXKTRILVTHSMSTLPQVDVITIV 840
QY 841 MSGKISIMSGSYOELLARDGAFAEFLRTYASTEOQDAENGVTGVSGPGKEAKOMENG 900
Db 841 MSGKISIMSGSYOELLARDGAFAEFLRTYASTEOQDAENGVTGVSGPGKEAKOMENG 900
QY 901 LVTSAGKQLOROLSSSSSYSGDISRHHNSTAELOKAEKKEETWKLMEADKAOTGVKL 960
Db 901 LVTSAGKQLOROLSSSSSYSGDISRHHNSTAELOKAEKKEETWKLMEADKAOTGVKL 960
QY 961 SVYWDYMKAI GLFISFLSIFLFCMCHVASALASNYWLSLMTDPIVNGTOEHTKYRLSYVG 1020
Db 961 SVYWDYMKAI GLFISFLSIFLFCMCHVASALASNYWLSLMTDPIVNGTOEHTKYRLSYVG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGGILASRCLHVDLHSLRSFMSFFERTPSGNLVNRFSEL 1080
Db 1021 ALGISOGIAVFGYSMAVSIIGGILASRCLHVDLHSLRSFMSFFERTPSGNLVNRFSEL 1080
QY 1081 DTVDSMTPEVIXMFMSGLFENVIGACIVILLATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
Db 1081 DTVDSMTPEVIXMFMSGLFENVIGACIVILLATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
QY 1141 KRLESVSRSPYVSHFNETLGVSVIRAFEOERFIHQSDLKVDENOKAYVPSIVANRWLA 1200
Db 1141 KRLESVSRSPYVSHFNETLGVSVIRAFEOERFIHQSDLKVDENOKAYVPSIVANRWLA 1200
QY 1201 VRLECVGNCTVLPALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSEMETNIVA 1260
Db 1201 VRLECVGNCTVLPALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMOIOETAPPSWPQVGRVEFRNYCLAYREDLDPVLRHINVTNGG 1320
Db 1261 VERLKEYSETEKEAPMOIOETAPPSWPQVGRVEFRNYCLAYREDLDPVLRHINVTNGG 1320
QY 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
Db 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGSIRMNULDPPSYSDERWTSLELAHLKDFVSALPDKLDHECAGEGENLSVGQROLVCL 1440
Db 1381 SGSIRMNULDPPSYSDERWTSLELAHLKDFVSALPDKLDHECAGEGENLSVGQROLVCL 1440
QY 1441 ARALLRRTKILIVLEATAVULEFTDDLIQSTIRTOFEDCTVLTAAHRLNTIMDYTRVYL 1500
Db 1441 ARALLRRTKILIVLEATAVULEFTDDLIQSTIRTOFEDCTVLTAAHRLNTIMDYTRVYL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531
Db 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531

Search completed: December 15, 2005, 15:17:10
Job time : 180.071 secs

(order) 1000 000 000

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:57:04 ; Search time 37.1218 Seconds

(without alignment)
4901.320 Million cell updates/sec

Title: US-10-665-283-4

Perfect score: 9734

Sequence: 1 MALRGFCSADGSDPLMDMNV.....NTIKVPTPLCTARQLDEDRS 1891

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database : PIR 80.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7849	80.6	1531	1	DVHNR
2	4484.5	46.1	1527	2	JP0336
3	3587	36.9	1545	1	S71841
4	3507.5	36.0	1541	1	S71839
5	3289	33.8	1494	2	E89447
6	3225	33.1	1573	2	T21219
7	3185.5	32.8	1502	2	T12216
8	2838	29.2	1515	1	S51863
9	2732	28.1	1398	2	T20434
10	2548	26.2	1478	2	T38712
11	2384.5	24.5	1623	2	T01359
12	2356.5	24.2	1622	2	DB6428
13	2326	23.3	1495	2	E86428
14	2312	23.8	1144	2	T27408
15	2294.5	23.6	1559	1	S64757
16	2263.5	23.0	1488	2	F86428
17	2238.5	23.0	1516	2	F84919
18	2230	22.9	1539	2	T78059
19	2144.5	22.0	1355	2	T00961
20	2144.5	22.0	1514	2	T52080
21	2140	22.0	1515	2	T52081
22	2131.5	21.9	1490	2	T7840
23	2129	21.9	1545	2	T46645
24	2123	21.8	1153	2	T26883
25	2109	21.7	1545	2	T42751
26	2104.5	21.6	1511	2	T42711
27	2098	21.6	1541	2	T42728
28	2069.5	21.3	1389	2	T47796
29	2062.5	21.2	1661	2	S64800

30	2037	20.9	1121	2	C87973	protein Y43F8C.12
31	2014.5	20.7	1582	2	A56248	sulfonylurea recep
32	2008	20.6	1592	2	A48933	probable transp
33	1888	19.4	1477	2	S64616	YORI protein - yea
34	1884.5	19.4	1037	2	T50518	ABC transporter-11
35	1838	18.9	390	2	S68403	inward rectifier P
36	1827	18.8	390	2	JC4689	inwardly rectifyin
37	1805	18.5	1146	2	F84487	probable ABC trans
38	1781	18.3	390	2	A57616	inward rectifier K
39	1770	18.2	390	2	JC7901	inwardly rectifyin
40	1722	17.7	946	1	JC5667	multidrug resistan
41	1693	17.4	1548	1	DVLNS	multidrug resistan
42	1687	17.3	1421	2	T34225	hypothetical prote
43	1606	16.5	1427	2	T20903	hypothetical prote
44	1575.5	16.2	1427	2	T39219	atp-binding casset
45	1558	16.0	1469	2	T50210	probable ABC trans

ALIGNMENTS

RESULT 1

DVHNR
multidrug resistance protein (cell line H69AR) - human
N/Alternate names: multidrug resistance-associated protein (MRP)
C/Species: Homo sapiens (man)
C/Date: 30-Jun-1993 #sequence_revision 05-Dec-1998 #text_change 19-Jan-2001
C/Accession: A44231, A37495
R/Cole, S.P.C.; Bhargava, G.; Gerlach, J.H.; Mackie, J.E.; Grant, C.E.; Almquist, K.C.;
Science 258, 1650-1654, 1992
A/Title: Overexpression of a transporter gene in a multidrug-resistant human lung cancer
A/Reference number: A44231, MUID:9308080, PMID:1360704
A/Accession: A44231
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 'MAPTSGTGMRGIPATPSPAPTRSSCGCVFTSGPV', 50-1531 <COI>
A/Cross-references: UNIPARC:UPI00001746CB; GB:105628; NID:91835658
A/Experimental source: small cell lung carcinoma cell line H69AR
A/Note: Sequence extracted from NCBI backbone (NCBI:119851); this sequence has been corrected
R/Cole, S.P.C.; Deely, R.G.
Science 260, 879, 1993
A/Title: Multidrug resistance-associated protein: sequence correction.
A/Reference number: A37495, MUID:93262415, PMID:8098549
A/Accession: A37495
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-60 <CO2>
A/Cross-references: UNIPARC:UPI00001746CC; GB:105628; NID:91835658
A/Note: sequence extracted from NCBI backbone (NCBI:131929)
C/Genetics:
A/Gene: GDB:MRP
A/Cross-references: GDB:136335; OMIM:158343
A/Map position: 16p13.1-16p13.1
C/Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C/Keywords: antibiotic resistance; ATP; duplication; nucleotide binding; P-loop; transmembrane
F/661-844/Domain: ATP-binding cassette homology <ABC1>
F/678-685/Region: nucleotide-binding motif A (P-loop)
F/1310-1503/Domain: ATP-binding cassette homology <ABC2>
F/1327-1333/Region: nucleotide-binding motif A (P-loop)
F/1450-1454/Region: nucleotide-binding motif B

Query Match 80.6%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	MALRGFCSADGSDPLMDMNVNTNTNTNDPFCQONTLVWVPCPYLMACPFYFLYSRH	60
DB	1	MALRGFCSADGSDPLMDMNVNTNTNTNDPFCQONTLVWVPCPYLMACPFYFLYSRH	60
QY	61	DRGYIQMTPLNKTALGFLMTVCVADLFYFWMERSRGIFLAPVFLVSPFLTGITTLA	120
DB	61	DRGYIQMTPLNKTALGFLMTVCVADLFYFWMERSRGIFLAPVFLVSPFLTGITTLA	120

```

QY 121 TFLIQLERKRGVSSGIMLTFMLVALVICALILRSKIMTALKEDAQVDLPFDITFYVFS 180
Db 121 TFLIQLERKRGVSSGIMLTFMLVALVICALILRSKIMTALKEDAQVDLPFDITFYVFS 180
QY 181 LLLIQLVLSGSDNSPLFSETHDPNCPDESSASLSTITWMITGLIVRGYROPLESD 240
Db 181 LLLIQLVLSGSDNSPLFSETHDPNCPDESSASLSTITWMITGLIVRGYROPLESD 240
QY 241 LMSLNKEDTSBOVAVLVKNNKKECAKTRKOPVKVYSSKDPAPOKESSKYDANEVEAL 300
Db 241 LMSLNKEDTSBOVAVLVKNNKKECAKTRKOPVKVYSSKDPAPOKESSKYDANEVEAL 300
QY 301 IVKSPOKEWNSPLFKVLYKTGFPYELMSFFPKAIHDLMMFSGPQILKLIKRVNDTKAPD 360
Db 301 IVKSPOKEWNSPLFKVLYKTGFPYELMSFFPKAIHDLMMFSGPQILKLIKRVNDTKAPD 360
QY 361 WQGFYFVLLFVYLAQTLVHLQYFHICFVSGMRIKTAVIGAVYRKALVITNSARKSTV 420
Db 361 WQGFYFVLLFVYLAQTLVHLQYFHICFVSGMRIKTAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLSVDARFMDLATYINMISAPLOVTLALYLMNLGSPVLAGVAVMVLMPVN 480
Db 421 GEIVNLSVDARFMDLATYINMISAPLOVTLALYLMNLGSPVLAGVAVMVLMPVN 480
QY 481 AVMAKTKTYOVANHKSKDNRIKLMEIILNGIKVLKLYAMBLAFKQVLAIROBELKVLK 540
Db 481 AVMAKTKTYOVANHKSKDNRIKLMEIILNGIKVLKLYAMBLAFKQVLAIROBELKVLK 540
QY 541 KSAYISANGFTWCTPPLVALCTPAVYVTTIDENNILDAQTRFVSLAFNLIRPDLNLP 600
Db 541 KSAYISANGFTWCTPPLVALCTPAVYVTTIDENNILDAQTRFVSLAFNLIRPDLNLP 600
QY 601 MVISIVASYSIKRLRIFLSHEELEPDSIERRPKDGGSNTIVRNATFTMASSDPT 660
Db 601 MVISIVASYSIKRLRIFLSHEELEPDSIERRPKDGGSNTIVRNATFTMASSDPT 660
QY 661 LINGITFISPEGALVAVGVGCGKSSLSALIAENDKYEGHVALKGSVAVYVPOQAMIOND 720
Db 661 LINGITFISPEGALVAVGVGCGKSSLSALIAENDKYEGHVALKGSVAVYVPOQAMIOND 720
QY 721 SLRENILFGCCOLEEYYSVIOACALLPDLLELPSGDRTIEGKGVNLSGGQKQVSLAR 780
Db 721 SLRENILFGCCOLEEYYSVIOACALLPDLLELPSGDRTIEGKGVNLSGGQKQVSLAR 780
QY 781 AVYSNADLYLPDDPLSAVDAAHGKHIFENVIGPKMLNKRIILVTHSMSTLPQVDVILV 840
Db 781 AVYSNADLYLPDDPLSAVDAAHGKHIFENVIGPKMLNKRIILVTHSMSTLPQVDVILV 840
QY 841 MSGKISMGSYOELLARDGAFAEFLRTYASTEOQDAEENGVTGSGGKEAKOMENGM 900
Db 841 MSGKISMGSYOELLARDGAFAEFLRTYASTEOQDAEENGVTGSGGKEAKOMENGM 900
QY 901 LVTSBAGQOLROQLSSSSSGDISGRHNSTAELOKAEKKEETWKLMEADAKQGYVL 960
Db 901 LVTSBAGQOLROQLSSSSSGDISGRHNSTAELOKAEKKEETWKLMEADAKQGYVL 960
QY 961 SVYMWYMAIGLFIPLSIFLPMCHVSLASNWMLSLMTDPIYNGQOEHKVLASYG 1020
Db 961 SVYMWYMAIGLFIPLSIFLPMCHVSLASNWMLSLMTDPIYNGQOEHKVLASYG 1020
QY 1021 ALGISQGLAVFGYSMAVSIIGLILASRCLVLDLHSLILSPMSFERTPSGNLVNFSKEL 1080
Db 1021 ALGISQGLAVFGYSMAVSIIGLILASRCLVLDLHSLILSPMSFERTPSGNLVNFSKEL 1080
QY 1081 DTVDMSIDBEVITKMGSLFNVIAGCTVILATPPIAIIIPRLGLIYFPVQRFYVASSROL 1140
Db 1081 DTVDMSIDBEVITKMGSLFNVIAGCTVILATPPIAIIIPRLGLIYFPVQRFYVASSROL 1140
QY 1141 KRLESVSSPYVSHNETLIGSVIRAFEOGRFTHQSLKXDNOKAYPSITVANRWLA 1200
Db 1141 KRLESVSSPYVSHNETLIGSVIRAFEOGRFTHQSLKXDNOKAYPSITVANRWLA 1200

```

```

QY 1201 VRLCEVNCIVLPAALFAVISRHSLSAGLVGSYSYLOVTTYINMLVRMSSEMETNIVA 1260
Db 1201 VRLCEVNCIVLPAALFAVISRHSLSAGLVGSYSYLOVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMGIOTAPSSWPQYGRVFRVYICRYRBDLPVLRHIVTTNGG 1320
Db 1261 VERLKEYSETEKAPMGIOTAPSSWPQYGRVFRVYICRYRBDLPVLRHIVTTNGG 1320
QY 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIILIPQDPVL 1380
Db 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIILIPQDPVL 1380
QY 1381 SGLIRMWLDPFSQYSDDEWVTSLELAHKDFVSALPKLHECAEGGENLSVGQROLVCL 1440
Db 1381 SGLIRMWLDPFSQYSDDEWVTSLELAHKDFVSALPKLHECAEGGENLSVGQROLVCL 1440
QY 1441 ARALLRRTKILVDEAIPAADLETDDLIOSTIRQFEDCTVLTARLNTIMDTRYVL 1500
Db 1441 ARALLRRTKILVDEAIPAADLETDDLIOSTIRQFEDCTVLTARLNTIMDTRYVL 1500
QY 1501 DKGEIOEYGAAPSDLLQORGLFYSMAKDAGLV 1531
Db 1501 DKGEIOEYGAAPSDLLQORGLFYSMAKDAGLV 1531

RESULT 2
JB0336
C:Species: Homo sapiens (man)
C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004
C:Accession: JB0336
R:Juchium, T.; Hinochita, E.; Haga, S.; Nakamura, T.; Tanaka, T.; Toh, S.; Furukawa, M.;
Biochem. Biophys. Res. Commun. 252, 103-110, 1998
A:Title: Isolation of a novel human canalicular multispecific organic anion transporter,
t.
A:Reference number: JB0336; MUID:99032812; PMID:9813153
A:Accession: JB0336
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-1527 <UCH>
A:Cross-references: UNIPROT:O15438; UNIPARC:UPI0000169859; GB:AF083552
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C:Keywords: ATP
F:1306-1499/Domain: ATP-binding cassette homology <ABC2>

Query Match 46.1%; Score 4484.5; DB 2; Length 1527;
Best local similarity 56.6%; Pred. No. 5.4e-295;
Matches 871; Conservative 261; Mismatches 372; Indels 35; Gaps 9;

QY 8 SADGSDPLMDNVMNTSNPDFTKCFONTVLVWVPCFYLMACPFYFLYSRRHGYIQM 67
Db 7 SGEIGSKRWDSNLSVHTENPDLPFCFONSILAWPCILYLMVALPCYLLYLRHRCGYIIL 66

QY 68 TPLNKTATGALFLIMYCMADLFYSFWERSGIFLAPVFLVSPULLGITTLATPLIOLE 127
Db 67 SHLSKLMVAGVLLMVCWMALEFYSFGLVGRAPAPVFFTPPLVVGVTMLATLILIOYE 126

QY 128 RRVQSSGIMLTFMLVALVICALILRSKIMTALKEDAQVDLPFDITFYVFSLLIQVY 187
Db 127 RLGQVSSGVLIITWFLCVCVAIVPFRSKILAAABEGISDPFFFTFYIHFAVLVSLAI 186

QY 188 LSCSDRSPLFSETHDPNCPDESSASFLSRTFMWITGLIVRGYROPLESDLSLNKE 247
Db 187 LACREKRPFFSAGVNDPNPERSAGLSLPLFWMTKMAIYGRHPLREKDLMSLKEE 246

QY 248 DTSQVAVLVKNNKKECAKTRKOPVKVYSSKDPAPOKESSKYDANEVEALIVKSPQK 307
Db 247 DRSQVAVVQQLLEAWRKQEKOTARH-----KASAAPGK---NASEDEVLLGAPRP 294

QY 308 EMNBSLFRVLYKTGFPYELMSFFPKAIHDLMMFSGPQILKLIKRVNDTKAPDMQGYRYT 367
Db 295 R-KDSFLKALLATGSSPLISACFKLIQDLISFNPOLSTILIRPISNPMGSPWMGFLVA 353

```

368 VLFEVTAQLQTLVHOFHILCFVSGMRITKIVAGVYKALVITNSARKSTVEIYVLM 427
354 GLMFLCSMMQSLIIQHYYHIFVTVGKERTIGMVIYKALVITNSVGRSASTVEIYVLM 413
428 SVDAQRFMDLATYVNMVMSAPLQVILALYILMLNIGPSVLAVGVAVLWVAVNVMAMKT 487
414 SVDAQRFMDLATYVNMVMSAPLQVILALYILMLNIGPSVLAVGVAVLWVAVNVMAMKT 473
488 KTVAVAHMKSMDNRIKLMNELLNGIKYKLYAMELAFMDKYLALROBELKYLKKSATLSA 547
474 RAFCVKMKDKSDRIKLMSEILNIGIKYKLYAMSEPSFKQYEGIRQGLQTLRTAAVLT 533
548 VGTFTWVCTPPLVALCTFAVVTIDENNIIIDAQAFVSLAFNLRPLNLPVIVSIV 607
534 TTTFTWMSPLVLTILMYVYVDPNVNDABKAFVSVSLFNLRPLNLPVIVSIV 593
608 QASVSLKRLIFLSHELEPDSIERRPVKGGSITVNRNATFMASSDPPLNGITFS 667
594 QASVSLKRIQFLSQBELDPQSVERTKISPG--YALTIHSGFTMAQDLPEPLHSLDIQ 650
668 IPEGALVAVGVQCCGKSSLSALLAEMDKVEGVAIKGSAVYVPOQAMIONDSLRNIT 727
651 VPKGALVAVGVQCCGKSSLSALLAEMDKVEGVAIKGSAVYVPOQAMIONDSLRNIT 710
728 FGCLLEPYSVSVIQAALPDLILPSSGDRTEIGKGVNLSSGOKORVSLARAVSNAD 787
711 FGKRLNKRKYQOTIEACALLADLEMLPGDOTEIGEKINISGOKORVSLARAVSNAD 770
788 IYFDDPLSAVDAAVHAKHIFENVIQPKGMLKNKTRILVTHSMYLPQVDVLIIVSGKIS 847
771 IFLDDPLSAVDAAVHAKHIFENVIQPKGMLKNKTRILVTHSMYLPQVDVLIIVSGKIS 830
848 EMGYSQELIARDGAFAPLRTYASTEOBQAEENGCVTVSGPGEAKOMENGM----- 901
831 EMGYSQELIARDGAFAPLRTYASTEOBQAEENGCVTVSGPGEAKOMENGM----- 890
902 ----VTSAGKQLOROUSSSSSSSGD-----ISRHNHSTA-LCKAEKKETMKMEAD 951
891 DNDVTVTVVQKQFNRQSLALSS-DGEGGRVPPRHLPSPSKVQVTEKKDGA--LTOBE 947
952 KAQGVKLVSVYDMYKAIQIFLSIFLPMCHNVSALSNVMSLMTDPIVNGTOEH 1011
948 KKAIGTELVSFMDYKAVAGLCTTLALCLVVGQSAALIGAVMLAMTMDAMDSQNN 1007
1012 TKRVLSTYGAIGISQGIATVGYMAVSTIGITLARSCHVDLHSLSPNSFFERTPSGN 1071
1008 TSIRLGYVAAIGIIQGFVLMMAAMAAAGGIIQAARVILHQAALHNKIRSPGFFPTPSGR 1067
1072 LVNRFSEKLDTPVDSMIEPVIKMFQSLFNVIGACIVILATPILAAIIIPGLGLYFFVOR 1131
1068 ILNCFSDKIYVDDVILAPVILMLNLSFNALSTLVIMASTPLTVVILPLAVLYTLVOR 1127
1132 FYVASSRQKRLAESVSRSPVYSHNETLLGAVVIRAEBOERFIHQSDLKVDENQAKYYP 1191
1128 FYAATSRQKRLAESVSRSPVYSHNETLLGAVVIRAEBOERFIHQSDLKVDENQAKYYP 1187
1192 SYVANRMLVRLBCVNCIVLFAALFAVISHRSISAGLVGIVSVSIVQVTTVIMLVMS 1251
1188 YIISNRMLVRLBCVNCIVLFAALFAVISHRSISAGLVGIVSVSIVQVTTVIMLVMS 1247
1252 SEMETNVAVERLKEYSETEKEAPWQIOETAPPSWQVGVGVEPRANCYARREDLVLR 1311
1248 SDESNSNVAVERVKEYSETEKEAPWQIOETAPPSWQVGVGVEPRANCYARREDLVLR 1307
1312 HINVTNGEKVGIIVGTAGKSSLTGLFRINESAGEIIIDGINAKIGLHDLARKKIT 1371
1308 DLSLHVGGEKVGIVGTAGKSSMTCLFRILAAAGEIRIDGLANVADIGLHVSQGLT 1367
1372 IIPDDPVLFGSGLAMNIDPQSYDEDEWVSLSLAHLKDFYSALPDKLHCECARGENLS 1431
1368 IIPDDPVLFGSGLAMNIDPQSYDEDEWVSLSLAHLKDFYSALPDKLHCECARGENLS 1427
1432 VGGQVLCLARALLRKTKIIVLDEATAVADLETDDLQSTIRTFQEDQVTLTIHRLNTI 1491

1428 VGGQVLCLARALLRKTKIIVLDEATAVADLETDDLQSTIRTFQEDQVTLTIHRLNTI 1487
1492 MDYTRVVLVDKGEIIOEYGAPSDILQORGLFYMAKADGL 1530
1488 MDYTRVVLVDKGEIIOEYGAPSDILQORGLFYMAKADGL 1526
RESULT 3
S71841
multidrug resistance protein, canalicular - human
C/Species: Homo sapiens (man)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C/Accession: S71841, S71840
R/Koenig, U.; Keppler, D.
Submitted to the EMBL Data Library, August 1996
A/Accession: S71841
A/Molecule type: mRNA
A/Residues: 1-1545 <KOE>
A/Cross-references: UNIPARC:UPI00001746CD; EMBL:X96395; NID:g1507819; PID:CA65259.1; F
R/Buechler, M.; Koenig, U.; Brom, M.; Kartenbeck, U.; Spring, H.; Horie, T.; Keppler, D
J. Biol. Chem. 271, 15091-15098, 1996
A/Title: cDNA cloning of the hepatocyte canalicular isoform of the multidrug resistance
A/Reference number: S71839; MUID:96279006; PMID:8662992
A/Accession: S71840
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1415-1429, 'VP', 1432-1455, 'E', 1457-1545 <BUE>
A/Cross-references: UNIPARC:UPI00001746CE; EMBL:X96395
C/Genetics:
A/Genes: GDB:ABCC2; CMOAT; ABC; MRP2; CMRP; DJS
A/Cross-references: GDB:6089489; OMIM:601107
A/Map position: 10q24-10q24
C/Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C/Keywords: ATP; glycoprotein; nucleotide binding; P-loop; transmembrane protein
F/654-837/Domain: ATP-binding cassette homology <ABC1>
F/671-678/Region: nucleotide-binding motif A (P-loop)
F/117-1510/Domain: ATP-binding cassette homology <ABC2>
F/1334-1341/Region: nucleotide-binding motif A (P-loop)

Query Match 36.94; Score 3587; DB 1; Length 1545;
Best Local Similarity 47.54; Pred. No. 3.7e-234;
Matches 746; Conservative 286; Mismatches 461; Indels 78; Gaps 20;
3 LRGFGSADGSDPLMDNVTWNTS-----NDFTKCFQNTLVVWPCFYLMACEPPYFL-- 55
2 LKFCN-----STFWNSSFLDSPRADLPFCFQGTLVVWIPGLFVLALPWQLAHV 51
56 YLSRHDRGIQMTPLANKTKTALGFLMIYVCMADLPFSFMRSGIFLAPVFLVSPITLGI 115
52 YKSRITKRSSTTKLYLAK-QVFGVFL-ILAAIILAVLITEDSOQAVPAVRVTPSLY-L 108
116 TLLATFLIQLERKQVQSGIMLT-FWLVALCALALRSKIMTALXEDAQVLDPRDT 174
109 GTWLVLILQYSQKQVQKSMPLSLFWIISILCGTFQFOFTLIRTLQGD-NSNLAVSCL 167
175 FYVYFSLILQVLSCFSDRSPLFSETIHDNPNCPSSASFLSRITFWMTGLIVGYRQ 234
168 FVISYQRIILILFSAFSEN-----NSSNNPSSIASFSLISYSWDSIILNGYKR 219
235 PLGSDLSMLNKEDTSBQVPLVYKMKKECAKTR-----OPVYVYSSKD 281
220 PLTLEVMEVDEEMKTKTLVSKETEMKRELQKARALQRRQKSSQONSQAGALPGILKN 279
282 PAQPKSSKVDANEVYALIVKSPKEMWNS--LFLVLYTGPFPVFLMSPFPAIHDLMM 339
280 QSGSQDALV---EDVEKKKKSGTKDDVPKSLMLALPFTFTWLKSLGLLVNDIF 336
340 FSGPQILKILIKFVNDTKADWQGYFYVLLFYTACLQTLVLHOFHILCFVSGMRITKAV 399
337 FVSPQILKILISFASDRDTLMIIGVLCATLLFPAALIQSFCLQCYFQLCFKLGVKVRTAI 396

Qy	400	IGAVYRKALVTN	SKRSTGEIYN	MSDAQRM	DLATYIN	TKSAPLOV	IILALYLM	455
Db	397	MASVYKKA	LVTN	LARKKEYT	GEITV	NMSVDAQ	KMDVNTMNM	456
Qy	460	LNIGSV	LAGVA	VWV	WV	PVNA	VWAMKRTKYOV	459
Db	457	RELGP	SV	LAGV	WV	LV	PIN	456
Qy	520	WELAF	KD	KVLA	ROBEL	KV	LKKSAYL	579
Db	517	WEPSP	RDO	VO	N	L	RKKE	576
Qy	580	QTAF	PS	LA	FN	IL	RP	639
Db	577	QKAF	PS	IT	LE	FN	IL	634
Qy	640	GTNS	IT	VNA	FT	WMA	RDP	699
Db	635	--KAM	Q	SE	AS	FT	W	692
Qy	700	GHVA	L	KGS	V	AAV	P	759
Db	693	GHIT	K	GT	TA	V	P	752
Qy	760	EIGE	GN	VL	SG	Q	K	819
Db	753	EIGE	GN	IL	SG	Q	K	812
Qy	820	KTRI	L	V	T	S	M	878
Db	813	KTRI	L	V	T	S	M	872
Qy	879	EENG	V	T	S	A	G	928
Db	873	-----	T	V	H	D	S	927
Qy	929	NSTA	E	L	Q	A	E	983
Db	928	RNSL	T	R	V	N	S	987
Qy	984	CNHV	A	L	A	S	N	1039
Db	988	MNSV	A	F	I	S	N	1047
Qy	1040	GGI	A	S	R	C	L	1099
Db	1048	GFVH	A	S	N	I	L	1107
Qy	1100	NVIG	A	C	I	V	I	1155
Db	1108	GI	S	T	L	I	C	1167
Qy	1160	LGV	S	T	R	A	P	1219
Db	1168	SG	L	P	T	R	A	1222
Qy	1220	IS	R	S	L	A	G	1279
Db	1228	IY	R	D	L	S	G	1286
Qy	1260	ET	A	P	P	S	S	1333
Db	1287	D	K	R	P	P	D	1346
Qy	1340	L	F	R	I	N	E	1399
Db	1347	L	F	R	I	N	E	1406
Qy	1400	W	T	S	L	E	I	1455
Db	1407	W	K	A	L	E	I	1466
Qy	1460	V	D	L	E	T	D	1515

Db 1467 VDLFTDNLQITTIQNEFHAHVITTLAHLHTTIDSDKMWLDNGKITIECGSPBEELQIPG 1526

Qy 1520 LFYSNAKDAGL 1530

Db 1527 PFYFAKEAGI 1537

RESULT 4

571839

canaliculular multidrug resistance protein - rat

C.Species: Rattus norvegicus (Norway rat)

C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C.Accession: S71839

R.Buchhalter, M.; Koenig, J.; Brom, M.; Kartenbeck, J.; Spring, H.; Horie, T.; Keppler, D.

J. Biol. Chem. 271, 15091-15098, 1996

A.Title: cDNA cloning of the hepatocyte canaliculular isoform of the multidrug resistance

A.Reference number: S71839; MUID:96279006; PMID:8662992

A.Accession: S71839

A.Status: preliminary; nucleic acid sequence not shown

A.Molecule type: mRNA

A.Residues: 1-1541 <BDB>

C.Cross-references: UNIPROT:Q63120; UNIPARC:UPI000012578; EMBL:X96393; NID:G1292881; PI

C.Superfamily: human multidrug resistance protein CMOAT2; ATP-binding cassette homology

C.Keywords: ATP, glycoprotein, nucleotide binding, P-loop, transmembrane protein

F.100-124/Domain: transmembrane #status predicted <TM01>

F.127-151/Domain: transmembrane #status predicted <TM02>

F.160-187/Domain: transmembrane #status predicted <TM03>

F.305-329/Domain: transmembrane #status predicted <TM04>

F.354-381/Domain: transmembrane #status predicted <TM05>

F.431-451/Domain: transmembrane #status predicted <TM06>

F.456-476/Domain: transmembrane #status predicted <TM07>

F.536-564/Domain: transmembrane #status predicted <TM08>

F.574-602/Domain: transmembrane #status predicted <TM09>

F.650-833/Domain: ATP-binding cassette homology <ABC1>

F.667-674/Region: nucleotide-binding motif A (P-loop)

F.966-994/Domain: transmembrane #status predicted <TM10>

F.1018-1046/Domain: transmembrane #status predicted <TM11>

F.1104-1132/Domain: transmembrane #status predicted <TM12>

F.1203-1228/Domain: transmembrane #status predicted <TM13>

F.1333-1506/Domain: ATP-binding cassette homology <ABC2>

F.1330-1337/Region: nucleotide-binding motif A (P-loop)

F.6,1007,1010,1011/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 36.0%; Score 3507.5; DB 1; Length 1541;

Best Local Similarity 46.2%; Pred. No. 9.1e-229;

Matches 721; Conservative 301; Mismatches 471; Indels 69; Gaps 21;

Qy 6 FCSADGSDPLDMWNTWMTSNPDTKCFONTVLVWVDFCLMACEPP--YFLYLSRRHRG 63

Db 4 FCNS----TFWPLSL--ESPEDLPLCFEQTVAWVPLGLFMLAPQLQSVSRSTRS 58

Qy 64 YQMTPLNTKALGLFWIWCADLFYSFWEBSRGFLPVPFLVSPFLIGITLLATFL 123

Db 59 SI--TFEYIAKQVFVFLILAILAIDLSLALTEBTGQATVPVRYTNP--ILYLCTWLVLVA 115

Qy 124 IQLERRKGVQSSGIMLT-FMLVALVQALAIRSKIMALEDAQVDPRDITFYVFSLL 182

Db 116 VQHSQMCCKRKNKSWFLSLFWILSVLCGVQFQI-LIRALKLSDKSNMAVSYLFFVSYFQ 174

Qy 183 LIQLVLSCFSDRSPLFSEFTHDNPCEBSASFLSRITFWITGLIVRGYQPLBGSPLW 242

Db 175 IYLLILITAFSGPS-----DSTQPTSVASFLSITFSFWYDRKVLKGYKHPFLLEBW 226

Qy 243 SLNK-----EDTSEQVVPVLVKN-----WKSCAKTRKQPVKVVYS-SKDPAQKESK 290

Db 227 DIDEGFKTRSVTSKEFAAMTKDLQKARQAFORLQKSORPEATLTHGLNKQSOQDVLV 286

Qy 291 VANEEVEALIVYSPOKEMNPSLFKVLKYTFGEYFLMSPEFKAIHIDLMFSGPOLIKLLI 350

Db 287 LBEAKKSKRTKDYPSK---LIRSLFKTFHVLILKSLFKLIDHLLVFLNPQLIKLLI 343

Qy 351 KFNVDIKADWQGYFYTVLLFVTACIQTLVLHGYFHCIVSGMRKIKTAVIGAVRKALVI 410

Db 344 GFVKSNSNYWVGYICALIMFAVTLIGSCLOSYQHCFVLMGCRITVWSSITYKALTL 403

QY 411 TNSARKSSTVGEIYNILMSVDAQRFYDLATYINIMISAPLOVILALYILMLNIGPSVLAV 470

Db 404 SMLARKQYITIGETVNLMSVDSQKMDATYMQVLWVMSVIGITLSIFFLMRRELGPISILAV 463

QY 471 AMVULMVPNPAVMAMKTKTYQVAHKKSKDNRIKLMNELLNGIKVLTALAMELAFDKYLA 530

Db 464 GVMVLLIPNGVLATKIRINIYQNNKNKDKRKIKNNEILSIGIKILKTFAMEPSFOEVOG 523

QY 531 IROBELKYLKSKAYISAVGTFTWCGPFLVALCTFAVYVTEIDENNILDAQTFVSLAFN 590

Db 524 IRKKELKULRQGQLOSLIFLQITPILVSVITSYVVLVDSANVLNKRKAFISITLFPN 583

QY 591 ILRFLPNLILPWNISIVQASVSLKRLRIFLSHEELPEPSIERRPVKDGGGTSITVRNAT 650

Db 584 ILRFLPNLILPWTSSILQASVSVDRLERYLGGDDDLDTAIRVSNFD----KAVFSEAS 639

QY 651 FTWASDPEPTLANGITFSPREGALVAVVGOGCGKSSLSALLAEMDKYBGAHAKGSVAY 710

Db 640 FTWDDLEBATIDVMDIKPGQTVAVVGVSCKSSLVSAMIGEMENYGHITIGGSTAY 699

QY 721 VPOQAMINDSRENIILFGCOLSEEPYRSVLOACMLDPLILPBGDRTEIGEKVNLISG 770

Db 700 VPOQSWIQNGITKNDILFGSEYNEKKYQOVLKACALLDPLEILPGDMAELGEEKINISG 759

QY 771 GOKQSVSLARAVSNADYILFDDPLSAVDAAHAKGKIFENVISPKMKMLKNKRTILVTHSMS 830

Db 760 GOKQSVSLARAVYQADYIILDDPLSAVDAAHAKGKIFKNVGPBNLAKGTRIFPTHGH 819

QY 831 YLPQVDYIIVNSGKISSEMGYSQELLABDGAELFKTYAS-TEQEQDAEENGVTGVSQP 889

Db 820 FLPGQVDEIYVLGKGTILKEGSTRYRDLDDKKGVPAPARWMTKFMKSGEGEATVNN----- 872

QY 890 GKEAKQEMNGMLVT-----DSAG-----KOLQOLSSSSSYSGDISRHNSTBELQ-- 935

Db 873 DSEADDDDDGLPYTEEIIPEDASLAMRENSLRTLRSSSSSSRRGSLKNSLTIKNV 932

QY 936 ---KAAKKEETWKLMEADKAQTQOVKLSVYMDYMKALGLFISFLIFLFCMNHVSALAS 992

Db 933 NVLKEKEKEVEBQKLIKKEPBTGKVKSIIYAKIYQANGWSILFILLFGLNNAVAFGS 992

QY 993 NYWLSLMT--DDEPVNGT--QEHTRKVLSSVYGALGISOGIAGVGSMAVSIIGILLASRCL 1048

Db 993 NLMISAMTSDSDNLNIGTNNSSSHRDMRIGVFGALGAGICILLISTISIYACRWASXAL 1052

QY 1049 HVDLHLSTLRSMSFEERTPGSGLVNRFSKELDTYDSMIPETIKMFGSLFNVIGACIY 1108

Db 1053 HGOLLTNILIRAPMRFDDTPTGRIVNRRSGDISYDULLPQTLRSMMWCFCFGIADTLVMI 1112

QY 1109 LLATPIAIIILPDLILFYFVOFRFVASSROKLRESYRSBVSVYHFMETLIGVIRAF 1168

Db 1113 CMATPVPALIIIPLSILYISQVFFVATSRQRLRLDSYTKSPIYSHSEETVGLTIIRAF 1172

QY 1169 EEOERFIHQSDLKVDENOKAYYPSIVANRWLAVRLECVGNCIVLPAALFAVISRHSLSAG 1228

Db 1173 EHOGQFLAMNEKQIDINQCVFSWITSNRMLAIRLEIGNLVFPCALLVIYKRTLIGD 1232

QY 1229 LVGLSVSLSQVTTYLANLVJRMSSMEETNIYAAVEKLKSYSEBKAQPMQIOETAPSSWP 1288

Db 1233 VVGFPLSALNALITQTLNMLVMTSEAFENIYAVERISIEYINVENEMAP-VYDKPPADWP 1291

QY 1289 QVGRVFPNVCJARVEDLDVFLRHNVTINGEKYGIQRTGAGKSSLTLLGFRINESA 1348

Db 1292 RHGEIOPNNYQVRRPELDVLKGTICNIKSEKGVGVRGAGKSSLTNCLFRILBSAG 1351

QY 1349 GEIILIDGINAKIGHDLRFKTIITIPQDVLFSGSLRMLNDFPQSYDSEBWTSLTALH 1408

Db 1352 GQIILIDGIDIVASIGLHDLRRLTIIIPQDPIFSGSLRMLNDFPNKXDEEYVRALELHL 1411

QY 1409 KDFYSALBDKLDHEAEGENLVSGORQVCLARLLAKTKIIVIDEKTAAVDLETDLI 1468

[illegible]

```

Query Match      33.8%; Score 3289; DB 2; Length 1499;
Best Local Similarity 47.6%; Pred. No. 5.6e-214;
Matches 696; Conservative 239; Mismatches 435; Indels 92; Gaps 18;

QY 119 LATELIQIERRRKGVSSGIMLTFWLVALVCALAIILRSKIMTALKEDEAQVDLFRDITTFVY 178
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 76 LALLITVACNKKGIITSGVITLWLVVCGIIPERFYLSGFIVNEVALBEIRATVYIIA 135
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 179 FSLLIQVLVSCFSDRBPFLFSETIHDNPCESSASPLSRITFWITGILVGRQPLEG 238
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 136 FTFSALBELFLCCPDA---VPSDMYKSSSSCCEYIASFINRLTFQWFGTAVLGNKKSDEN 192
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 239 SDLSLNKEPDSQGVVPLVGNMKECAKTRKQGVKVVYSSKDP--AQPKSSKVDANEE 296
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 193 EDLMDINEIDKAENLISPFQNLKPRIDEXH-QNIK-----KDPSSALPKN----- 237
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 297 VEAIIVSPQKENNPSLFKVLKTFEGPYELMSPEFFKAHIDLMSSGOILKLIKFNVDY 356
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 238 -----HPSFVPIPIFKYKYTLAAGFYKLCFDMLOFLAQLQLQLGLGFEDK 284
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 357 KAPDMOGFYTVLLHEFVACLOTLVHIOYFHI-CFSGMRKIKAVIAGVYRKALVITNSARK 416
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 285 NQPMWIGSIVIGMIFPSSFLQSMFLHQYHSMFLGMHGVSVLTSAVYSKALNINSNEARK 344
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 417 SSTGVEIYNLMSVDAQRFMDLATYINMIWSAPLOVLIATLYLLMLNLPVLAGVAVVLM 476
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 345 GKTIGALYNLMSVDIQIKIDMAPITMLFWSAPLOIFLSTYFLWFLGVAALAGLVILIA 404
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 477 VPVANAAMKTKTYQVAHMSKDNRIKLMNEILNGIVLKLIVAMELAFKQVLAIROEEL 536
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 405 LPVNGLIATIQRRKQCTQMKLKDERIKMGEIILNGMKVILKLYSVERSMENNVLKIRREL 464
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 537 KVLKKSAYLSAVGFTFWVCTPPELVALCTPAVYVYIID--ENNILDAQTAFVSLATNIIIRFP 595
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 465 HILKKSIFYMAAIVFSWICAPFLASVLSFVVVYVLDENNVLDELITFVALSTLDPILRMP 524
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 596 LNIKPMYISSIVQASVSLKRLRIPLSHEELEPDISERRPVQDGGTUSITVRNATFTPMAR 655
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 525 LAMVAMVYGAIVQCSVSNTRLKEFFLADEMSPOI-----SISHGETDAIIVENGILFWSMS 580
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

Qy	656	ADPPLNIGTISREPGALVAANVGVCCKSSLSLLAAMDVKEGVALIKGSAAVYPOOA	715
Db	581	DEDPRLREISFEKIQKQALVATKGKSSLSLHALLGEMNKLSGVQNGNAAVYPOOA	640
Qy	716	WIONDSLENTILFGCOLBEEPYRSVLOACALLPDLIELPESGRTIEIGEKGVNLSGQOKR	775
Db	641	WIONDSLENTILFNNRPYDLENEDVKNCALMEDLANLPAQRTIELGEGKINLSGQOKR	700
Qy	776	VSLARAVYSNADIVYFPDPLSAVDHAVHKGHIFENVIGPK-GMLKNKTRILVTHSMSYLPQ	834
Db	701	VSLARAVYQNPDIILLDPLSAVDSHVHGKIEFENVISSSTGCIASKTRVLYTHGLTYLKH	760
Qy	835	VDVILVMSGGKISEMSVOELLAROSAFAEPLRTYASTQEOBDAEENGTVGVSFPG----	890
Db	761	CDQLVTLKEGTISELGTYOELLNNSGAFAPLEEFILIESKTRGRVAYSIGDGSGEVDEIL	820
Qy	891	KEAKOMENGMLVTDGAKOLOROLSSSS-----SYSGDISRH---HNSSTLEOK	936
Db	821	RDLOGVKRGIL-----KRLESHLIGESDKEPDSARAIERYSRSSRSRVLHSPRGQHE	874
Qy	937	AEA-----KKEBTWKLMEADKAQTQGVKLSTYTWYMKAIGLFISPLSTLFMCNHY	987
Db	875	NEALLGAISEDPVPAOENTOLIEKETVETKXVFENVIAFYOAISIPITLLFFPLYGSSG	934
Qy	988	SALASNNYMLSLMTDD-PIVNGTOEHKRYLVSYGALGISOGIAVPEGSAVSIIGGLASR	104
Db	935	LGISLNFYIAKLSDHAKSGNRTSSDAKMLGIAVYAGQGSFVLLIASITLLTGVLASR	994
Qy	1047	CLHVDLHSLILSPMSFFERTPESGNLVNRFSEKLDYVDSMIPEVIMFMGSLFNVIGACI	110
Db	995	ILHAGLLGNIMHSPMAFPVTPILGRLLNKGDIEMIDRLPVLINHMSTINNVATLY	105
Qy	1107	VILLATPIAAIIIPPLGLIYFVQRFYVASSROKLKLESVSRSPVYSHENETLGVSVIR	116
Db	1055	VIMATPMPWAGIFAFALISVYIFVLRYSITSRQKRLLESASRSPISYHFOESIOGASAIR	111
Qy	1167	APEBERFHOGLKVDENQKYYPSBIVANRMVLALEECNGCIYFALAFNAYSRS--	122
Db	1115	AFGVVDNFIKSQOQRVDDHLIAYPSBIVANRMVLALEMGNLVLISAGAAYFPDPSG	117
Qy	1225	LSAGLVGHSVSYSLQVTTYLANMLVRMSSEMETNIVAVERLKESETEKEAPMOIOETAPP	128
Db	1175	LSAGLVGHSVSYSLQVTTYLANMLVRMSSEMETNIVAVERLKESETEKEAPMOIOETAPP	123
Qy	1285	SSWPOVGRVEFPNNYCLIRYEDDLFVLRHINVTINGEKVIGVGTAGKSSLTGLPRIN	134
Db	1233	KSWPKEGEISINFSVRYRPGDLVLHGISAHTAPEKVGIVGRTAGKSSLTGLAFRII	129
Qy	1345	BSABEIIIIIDGINIKIGHLDLRFKTTIIPQDVULFSGSLRMULDPEFSQSDSEWTSLE	140
Db	1293	EADGSGIEIDGINIANMLQLEQLSCLTIIYQDVLVLSGTWKNMLDPSAISDSQVWEALB	135
Qy	1405	LAHLKDPVSALPDKLDHECABEGENL-----SVGQROLVCLARALLRKTKI	145
Db	1353	NAHLKPFVKSLODGLFHKISSEGENNLRPAGSKSQKQQLNFSVQOROLICLARULLRKTIV	141
Qy	1451	LVLDATAAVALDETBDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVILDKGEIOEYGA	151
Db	1413	LVLDATAAVALDETBDLIQSTIRTOFEDCTVLTIAHRLNTIMVDSDRLLVLDKGRAVAEPDS	147
Qy	1511	PSDDL-QQKGLFYMAKDAQVY	151
Db	1473	PNULLANPDGIFYMAKDAQVY	1494

RESULT 6
T21219
hypothetical protein F21G4.2 - *Caenorhabditis elegans*
C|Species: *Caenorhabditis elegans*
C|Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #next_change 09-Jul-2004
C|Accession: T21219, T24002
R|Mortimore, B.
submitted to the EMBL Data Library, October 1996

A:Reference number: Z19392
A:Accession: T21219
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1573 <WT>
A:Cross-references: UNIPROT:O93552; UNIPARC:UPI00000823F7; EMBL:Z81016; PIDDN:CAB02667.1
A:Experimental source: clone F21G4
R:Kershaw, J.
submitted to the EMBL Data Library, March 1995
A:Reference number: Z19828
A:Accession: T24002
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1573 <WT2>
A:Cross-references: UNIPARC:UPI00000823F7; EMBL:Z48621; PIDDN:CAA88549.1; GSPDB:GMO0028;
A:Experimental source: clone R07B1
C:Genetics:
A:Gene: CESP:F21G4.2
A:Map position: X
A:Introns: 21/3; 122/3; 197/3; 276/1; 654/1; 911/1; 1067/1; 1472/2; 1537/2
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

```

0Y 3 LRG---FCSAGSDJLMMNMYWNTNSN--PDEKFCQONTLVWVPCFYLMACPEFYLYLS 58
Db 9 LKGDIAFC-----GGEYMD--PAMWNSYPIILSQCYOHTLVWF----- 47
0Y 59 RHDRGYIOMTPLNKTALGFLMLIVCADLFYSF-----WER-----SR 98
Db 48 -----TAIVFLDALITLQAIYFRRRPNPIWTRRIQLKIGLACILIAIDL 91
0Y 99 GIFIAPVVF-----LVSPILGITTLLATFLIQLERBKQVOSSGIMLTFVLVL 146
Db 92 SLFTAIETLEFQGPYAVDFPYPLTLCIAMVYLLIVSCNNGYIVSGLFISSLVFT 151
0Y 147 VCAALILR---SKIMTALKEDAQVDLFEDITEFYVYFSSLLIQLVLSCFSDRSPL--FSETI 202
Db 152 ISAIPELLYMIQOIVNPAEMAMWIDYPRCIAFIWFPCAFETYHACYADASPEGYKYL 211
0Y 203 HDNPNCEPSSASFLSRITFMMWITGLIVNGYRPLEGSDJLSLNKEDTSQOVYVPLVKNK 262
Db 212 AARNSPETTSFLNRITMMWPNFNSICSGVKKPLEVSDLYSNEADTSLMLVPKMYNLD 271
0Y 263 KECATRRKOPVNVVYSSKDPQPKSSKYDANEVEAL-----IYKSPQKWN 310
Db 272 KQSKKFEETPAARRIGSN--ASRTNRRRRTSSNDTTLPLNDQSTDDYGSVPAGOSTX--M 327
0Y 311 PSLFVLYKTFGPPYFLMSEFFFAIHDLMFSGPQILKLIKRVNDTKADWQCYFTVLL 370
Db 328 PSIIITLFLMFEMWDITAMFVLLSDVLLFCNPLBLKLSIRFTEEBERPMOGVVLATFM 387
0Y 371 FYTACLOTLYVHQYHICFVSGMRLKTAIVIGVYKRALYITNSARKSFTVGEIVNMSVD 430
Db 388 FFSALSTSLLSHYLYLMYLRVSTRJOTCLTAAYVYKTLRLSNARKREKTVGEIVNMAID 447
0Y 431 AORFMDLATYINMWSAPLOVTLIAYLLMLNLGFSVLAVGAVAVMLVPVNAVAMKTKTY 490
Db 448 VDRFOQITPQTMQYNSNPFQIOLATFLFLFQOLGVSFSGVAAMVLLFPINFIYTWIIRKM 507
0Y 491 QVAHMKSDNRIKLMNEILINGIKULKIYAMELAFKDKVLAIOEBELXUKSAYLSAVGT 550
Db 508 QIAQYMYDERKXMYNEVLNGIKVILKYAMEBPMQOVLEDREKQELGIKKAAPLETSED 567
0Y 551 FTWVCYFPLVATCEAVVYITDENNIIIDPAQTFVSLATFNIRFPLNIIIPMYISSIOAS 610
Db 568 MLNTASPLVIALSTFATFIYIDPKVILPEIAFVSLTFENQORSBMSQYABELITQVVOYV 627
0Y 611 VSLKRLRIFLSHEEIEPDSIERRPYKXGGSNTSIVRAKATFW--ARSPD--PLNGIITS 667
Db 628 VSNRLKEVLSEBELNEAIDHR--ARD--NDNDVICKEKCLSMESGSEHQPVPPLTNISFS 684

```


Db 887 -RRPPSDAAPVKS-----TSEAQMEPSLDDVEVTGLTAGEDSDVOYGRVKATYLSYIRA 940
Qy 970 IGLISFLSIFLFMCNHYASALASNYMLSLMTDIPVNGTOEHTKRLSVYALGISQGIA 1029
Db 941 VGTFLCTYTLTFLFCQYVASFCQGYWLSLMDDPVDDKQMSALRGSIFGLGLQAG 1000
Qy 1030 VFGYSMAVSIIGIILASRCLHVDLHLSILRSPMSFEERTPSGNLVNRFSEKLDYDSMPE 1089
Db 1001 LFASMAAFGLGARASCLFRSLMDVARSPIGFERPVGNLNRFSEKTDIVDVDPD 1060
Qy 1090 VIKRMGSLFENVIGACIYIILLATPIAIIIPPLGIYFVGQRYVASSRQLKRLSEVRS 1149
Db 1061 KMRLLTYAFGLLEVGLAVSMATPLAIVAILPLMLLYAGFOSLYVATCCQLRLESAYS 1120
Qy 1150 PYSHFNFTLLGVSYIRAFEEQERFIHOSDLKVDENOKAYPSIVANFMLVRLCEVNC 1209
Db 1121 SVCSHLAETFGSQYVRAFOAGPPTAQHDALMDENQGISPPRLVADKMLANLELLGNG 1180
Qy 1210 IVLPAALFAVISRHSLSAGVLGVSYSLQVTTYINMLVRMSSEMETNIIVAVERLKEYSE 1269
Db 1181 LVFVAATCAVLTKAHLISAGLAGFSVSAALQVOTQLQMTVRSWTDLENSMVAVERVQYVH 1240
Qy 1270 TEKAPMIOETAPSSSPQVGRVFRVYCLRYREDLDFVLHIVTNGEKVIGVIRT 1329
Db 1241 TPKEAPMFLPSSAOPMPCCGQIEFRDPLHREPLPMVAVGVSLKTHAGEKGVIGRT 1300
Qy 1330 GAGKSSLTGLPRINSAAGEIILIDGINIAKIGLDLHFKITIIPODPVLFGSGLRMND 1389
Db 1301 GAGKSSLTWGLLRLOAEAREGGIWDGVPITDMGLHTLASRTIIIPQDVLFPGLSRMND 1360
Qy 1390 PFSQYSDERVWTSLELAHKDFVSALPDKLDECAEGGBNISVGRQIVCLARALLRTK 1449
Db 1361 LQENTWDGIMAALETYVCLKAFVTSIPGLQYECGGQDLSVCGKOLLCLARALLRTKQ 1420
Qy 1450 ILVDEAFAAVALDETDLLQSTIRQFEDCTYTLTAHLNLTMDTRYIVLVDKGEIOYRG 1509
Db 1421 ILIIDEAFAVDPGTEIGMOALERWFAQCTVLLIAHRLRSNCARVLDDEGVABESG 1480
Qy 1510 APSDLLQOQGLFYSMAKDAGL 1530
Db 1481 SPAQLLAQKGLFRLAQSGL 1501

RESULT 8
S51863
cadmium resistance protein YCF1 - Yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YP9302.11c; protein YDR135C
C:Species: Saccharomyces cerevisiae
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S51863; A55352; S50233
R:Oliver, K.; Harris, D.
submitted to the EMBL Data Library, February 1995
A:Reference number: S51853
A:Accession: S51863
A:Molecule type: DNA
A:Residues: 1-1515 <OI1>
A:Cross-references: UNIPROT:P39109; UNIPARC:UPI000013A67B; EMBL:Z48179; NID:g665657; PID:R-Szczypka, M.S.; Wemmie, J.A.; Mowse-Rowley, W.S.; Thiele, D.J.
J:BIOL. Chem. 269, 22853-22857, 1994
A:Title: A yeast metal resistance protein similar to human cystic fibrosis transmembrane A:Reference number: A55352; MUID:94357936; PMID:7521334
A:Accession: A55352
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-679,'R',681-1515 <SZC>
A:Cross-references: UNIPARC:UPI0000168F06; GB:L35237; NID:g556464; PIDN:AAA50353.1; PID:C:Genetics:
A:Gene: SGD:YCF1; MIPS:YDR135C
A:Cross-references: SGD:S0002542; MIPS:YDR135C
A:Map position: 4R
C:Function:
A:Description: required for cadmium resistance
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

C:Keywords: ATP; nucleotide binding; P-loop; transmembrane protein; yeast vacuole
F:287-308/Domain: transmembrane #status predicted <TM>
F:345-366/Domain: transmembrane #status predicted <TM>
F:421-442/Domain: transmembrane #status predicted <TM>
F:446-467/Domain: transmembrane #status predicted <TM>
F:533-555/Domain: transmembrane #status predicted <TM>
F:558-580/Domain: transmembrane #status predicted <TM>
F:564-829/Domain: transmembrane #status predicted <TM>
F:663-670/Region: nucleotide-binding motif A (P-loop)
F:951-972/Domain: transmembrane #status predicted <TM>
F:995-1016/Domain: transmembrane #status predicted <TM>
F:1068-1088/Domain: transmembrane #status predicted <TM>
F:11092-1113/Domain: transmembrane #status predicted <TM10>
F:1179-1200/Domain: transmembrane #status predicted <TM11>
F:1208-1229/Domain: transmembrane #status predicted <TM12>
F:1289-1483/Domain: ATP-binding cassette homology <ABC>
F:1306-1313/Region: nucleotide-binding motif A (P-loop)

Query Match 29.2%; Score 2838; DB 1; Length 1515;
Best Local Similarity 40.3%; Pred. No. 2,Je-183;
Matches 632; Conservative 295; Mismatches 495; Indels 146; Gaps 27;

Qy 28 DETKCFQNTVLVWVPCFYLMACPFYF-----LYLSRH-----RGYIOMPTLANKTKT 75
Db 27 DFTQCFIDGVTL-----NLSAIFMTTIGIRDVLNLCCKKSGIKYRRRWITVSMALVTL 81
Qy 76 ALGFLMLVWCADLFYFEMERSRGIFLAPVFLVSPITLIGITTLATFLQLERRKGVQS 135
Db 82 EIAF-----VSLASINISKEAEN---FTYVQVASTLSLFLVALAHMIEYDR--SVYAN 132
Qy 136 GIMTFMWLVALVCLALIRSKIMTALKEDAQVDLFRDITFVYRS-----LLIQVLSC 190
Db 133 TVLFWYMLFETPGNFAKLINILIRHTYEG-----IWSGQGFILTLFQYITC 180
Qy 191 FS-----DSPL-----FSETIHDNPPCESSASFLSRTIPMWTGLIYRGYROPLEG 238
Db 181 ASILLLEALPKKPLMHQIHQTLIRKKNPDYSDANSFSRTFSMSGSLMTGETKYLVE 240
Qy 239 SDLMSLNKEDTSQVQVPLVYVNMKKCECAKTRQPKQVYVYSSKDAOPKSSSKVANEVE 298
Db 241 ADVLKLPNPFSSSELSQLEKNMENL-----KQKS----- 271
Qy 299 ALIVKSPQKWNPSLFKVLVYTFGPYFLMSFPFPAIHDLMMFSGDILKLIKTVND-- 355
Db 272 -----NPSLSMAICRTFGSKMLAFAFKAIDHVALFTOPOLRIILIKETVTDYNS 320
Qy 356 -----TKAPDMQGYFTYVLLFTYACIQTLVHOYFHICVSGMRITKA 398
Db 321 ERQDDHSSLQGFENNHPQKLEIVRGFLAFAFMPLVGFTQTSVLHQYFLANVTNTGYIKSA 380
Qy 399 VIGAVYRKALVITNSARKSTVGEIVNLMVSDAOEFMDLATYINMINSAPLOVILATYLL 458
Db 381 LTAALIQKSLVSNSEASLSTGDIIVNLMSTDVQKQDLTQWMLNLSNGPQIIICLYSL 440
Qy 459 WNLGSPVLAVAVVAVVAVPVNAVAMAKTKTYQVAHMKSKONRIKLANEILNGIKVLY 518
Db 441 KYLGNSMWGVITILVIMPLNSFLMRLOKLOQSKQMKQKDERTRVSEILNINIKSLKY 500
Qy 519 AMELAFKQKVLAIR-QEELKVLKKSAYLSAAGTITWCTPELVNLCFPAVVTIDENNIL 577
Db 501 AMEKYRKRLKEVNNKELKVLTKQCYMAATYSQFNVPVLPSCFFPAVTV-YTEDAL 559
Qy 578 DAQFAVSLALFNILRPLNLTLPWVSSIVQASVSLKRLRIFLSHEELPEPSIERRPVKD 637
Db 560 TTDLVFLPLTIPNLISFLPLMIIPVNLNFIASVISIGLFTFTNBELOPDSVQRLPKVK 619
Qy 638 GGGTNSITV-RNAITFTWASD--PPTLNGITFSIPEGALVAVVGQVGGKSSLSALLAE 694
Db 620 NIGDVAIINIGDAFTFLMRKREYKVALKNINFOAKKGNLTICIVGRVSGKTAALLSCMGD 679
Qy 695 MDKVEGHAIVKGSVAAYVQQAQMIQNSLRNLTIRGQLEBEYYSVYQAQALLPLDELTP 754
Db 680 LFRVVGFAIVGSAVAYVQVPMNGVTKENILFGHRVDAEFYKTIKACALTTDLAILM 739

[illegible][illegible]

A:Residues: 1-255 <BOY>
A:Cross-references: UNIPARC:UPI00001746CF
A>Note: the authors translated the codon CAG for residue 248 as His
R:Purnelle, B.; Goffeau, A.
Submitted to the EMBL Data Library, April 1996
A:Description: The sequence of 32 kb on the left arm of yeast chromosome XII reveals 14
mily and a new ABC transporter homologous to the human multidrug resistance protein.
A:Reference number: 569380
A:Accession: S69391
A:Molecule type: DNA
A:Residues: 1-1559 <PUB>
A:Cross-references: UNIPARC:UPI0000126A9C; EMBL:X97560; NID:G1297003; PIDN:CAA6162.1; F
R:Mosga, T.; Zimmermann, F.K.
Yeast 12, 693-708, 1996
A:Title: Sequence analysis of the CEN12 region of Saccharomyces cerevisiae on a 43.7 kb
e conductance regulator protein CTR.
A:Reference number: 570557; MUID:96405918; PMID:8810043
A:Accession: S70560
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1559 <MTW>
A:Cross-references: UNIPARC:UPI0000126A9C; EMBL:X91488; NID:G1495203; PIDN:CAA62776.1; F
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1995
C:Genetics:
A:Gene: SGD:BPT1; MIPS:YIL015W
A:Cross-references: SGD:S0003938
A:Map position: 12L
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C:Keywords: ATP; nucleotide binding; P-loop; transmembrane protein
F:28-44/Domain: transmembrane #status predicted <TM1>
F:143-159/Domain: transmembrane #status predicted <TM2>
F:178-194/Domain: transmembrane #status predicted <TM3>
F:334-350/Domain: transmembrane #status predicted <TM4>
F:421-437/Domain: transmembrane #status predicted <TM5>
F:526-542/Domain: transmembrane #status predicted <TM6>
F:550-566/Domain: transmembrane #status predicted <TM7>
F:654-679/Domain: ATP-binding cassette homology <ABC1>
F:672-679/Region: nucleotide-binding motif A (P-loop)
F:974-990/Domain: transmembrane #status predicted <TM8>
F:1017-1033/Domain: transmembrane #status predicted <TM9>
F:1099-1115/Domain: transmembrane #status predicted <TM10>
F:1118-1134/Domain: transmembrane #status predicted <TM11>
F:1212-1228/Domain: transmembrane #status predicted <TM12>
F:1319-1559/Domain: ATP-binding cassette homology <ABC2>
F:1336-1343/Region: nucleotide-binding motif A (P-loop)
Query Match 23.6%; Score 2294.5; DB 1; Length 1559;
Best local similarity 35.1%; Pred. No. 1.5e-146;
Matches 574; Conservative 292; Mismatches 541; Indels 227; Gaps 39;
QY 23 NTSNPDPTKCFONTVLVWVCFYIM-ACFPFYLYLSRHDRGYIQMTPLNKTATAGFL-80
DB 23 NALNP-----CFISVISAQAVFLLIGSYOLMKLYKNNKVPFRKKNPPTLPKINSRHLLT-78
QY 81 -LWIVCW-----ADLFYSPWBSRGIFLAPVLPVSPVLLGI-----TTLTAT-121
DB 79 HLTNVCQSTLIICELALVYQSSDRVPTLLK-KALTLNLF-----NLGISLPTQYLA-132
QY 122 FLIQLERRKVOSSGIMLTFMLVALYCALAIRSKI-----MTALKEDAOV-167
DB 133 F-----KSTFGMGNQLFYVMFOILLQFLILORYHSSNRLTYISQGTMIILEVL-185
QY 168 DLFRDITFYVYFSLLIQLVLSCPSDRPLPSETIHDNDCPSSASFLSRITFWITGL-227
DB 186 -LNSVAIFLY-DLCIFEPT-----NELSEYKKNQGWPF-----PVHVLSTYITFIWMNKL-233
QY 228 IYVGYROPLESGDLMSLNLK--EDTSEGVVPLVAVNMKKECAKTRKQGVKVVYSSKQDAP-285
DB 234 IYVETFR-----NKIKIPNOLPLPVDLNLK-----259
QY 286 KSSSKVDANEVEVALLVYKSPQKEMNPSLFLVLYKTFGPYLMGFPPKAIHDLMMFSGPQI-345
DB 260 SISKEPKANWELKWLNRN-----SLMRAIWKSPFRTISVAMLYETTSDDLSSVQPPQ-312

QY 346 LKLIRKFN--DTRAPDMQGYFYVTLFYTACQLVLMHOYHICFVSGMRITAVIGA-402
DB 313 LRIFIDGLNBPSTSSKPLPLNGVFTALFLVIVSVVFLTQPIGIRGAGIGIRGSLAST-372
QY 403 VYKALVITNSARKSSTVGEIVLMSVDA--ORFMDLATYIMMWAPQVILATLWL-459
DB 373 VYQKSLRLTLAENKSTGDIWLMSDVLRIGRFFENA--QTIIQAPQIIIVLVLSLV-429
QY 460 LNIQPSVLAQVAVMTLVVAVNVAVMKTTYQVAHMSKNNRIKLNELINGIKVLKYA-519
DB 430 WLIGKAVIGGLVMTALMMPINAFSLRKYKSLSTQMKYKDMRIKITELNAAIKSLYA-489
QY 520 MELAFQKVALAQE-ELKYLKKSAYLSANGTWTCTPPLVALCTEAVVITDENNILD-578
DB 490 WEPMARLHVHNDKELKFRKIGIVSNLIYFAMNCVPLMYCTSTGLP-SLPSDPLS-548
QY 579 AQTAFFVSLAFNLIRPLNLIPMWISISIVQASVSLKRLIFLSHELEPDSIER-RPVKD-637
DB 549 PALVFBLSLPLNLSAIVSPSMINTIITSVMERLKSFLSDEIDSFIERIDPSAD-608
QY 638 GGGTNSITVNRATFTW-----ARSDPPTLNGIT-----FSIPEGALV-674
DB 609 ERALPALNMNNTIFLMKSKVELTSSQSGDNLRTDEESIIQSSQIALKNIDHFEAKRGDLY-668
QY 675 AVYGVQCGCKSLLSALLAEMDKVEGH-----VAIKSVAYVPOQAMTQNDSLRENT-726
DB 669 CVVGRVAGKSTFLKAILGQLPCMSGRDSIPKLLIRSSSVAYVCSQESVIMWASVRENI-728
QY 727 LFGQLEPEPFRYSRVIOACALLPLEILPSGDRTEIGKGVNLGGGQKORVSLARAVSNA-786
DB 729 LFGHKRPQDYDITIRACQLLPDLKILPDGDELTVGKGISLGGQKARSLARAVSRA-788
QY 787 DIYLPDPLSAVDAAVHAKHIFENV-IGPKMLKNKTRILVTHSMYSYLPQVDVITVNSGK-845
DB 789 DITLDDIISAVDAEVSKNILEVVLGKTKALNKNKTLITLTVNVSILKSKQMTYALENGE-848
QY 846 ISMGYSQIELAR--DGAFAEFLRTYAS-----TEQEDDAEENGVTGSGPGKEAKOME-897
DB 849 IYVQGVYEDVMNKNNTSKLKLLEBFSDPIDNGNSDVQTEHRSSEBDEP-----900
QY 898 NGMLVYDSAGKOLROLOLSSSSYSQDISRHNSHTAELOKAEAKKEFWKL-----947
DB 901 -----IQLVKTBSETEDEAVT--BSELEIRKNSRSLATLRPPFVGAQL-945
QY 948 -----MEADKAQQTQGVKLVYVWDYMKAIIGLIFSPISIFLFW-CHNVSAIASNYLSLW-999
DB 946 DSVKTAQKAEKTEVGRVTKIYLAIVIKCGV-LGVVLPFLFMILTRVPDLANFPLKYM-1004
QY 1000 TDDPIYNGQEHTRKRLSYGALGISQGLAVFGYSAVSI--GGILASRCLHVDLLHSI-1056
DB 1005 SESENERKSGNERVWMEVGYSLIGVAS--AAFNLLRSIMMLLYCSIRGSKJHESNAKSV-1062
QY 1057 LRSPEMFPERTEPGNIVNRPSEKLDIVDSIMPIEVIMFGSLPNVIGACVILLAPPIA-1116
DB 1063 IRSMTFFETTPGRILINRFSDDMDAVDSLQYIFSPFKSILFYIVTYVLVGYNMPWL-1122
QY 1117 IILPGLIYFFVQRYVYVASSRQLKRLAESYSRSRPVYSHFNETLGVSVIRAFEEQERFI-1176
DB 1123 VFMPLVLIYIYQGTIYVLSRELKRLISISYSPINSLMSSESLNGSIIDAYHFRFPIY-1182
QY 1177 QSDLKVDENOKAYPSIVANRWLAFLVCEVGNCTVLPALFAVI--SRHSASGLAVGSL-1233
DB 1183 LNYEKIQYVNDVFEVFNERSTRNWSVLQGTIGATIVATATAILATMNTKQSLSSGMVGL-1242
QY 1234 VSYSLQVTTYLNMVLMSSSEMETNIVAVERLKEYSYTEKBAPOQIOTLPSPSPPOGVY-1293
DB 1243 MSYSLEVTGSLTYIATYTYITINIVSVRIYVEYGLPPELAGSINERKPDENWPSKGI-1302
QY 1294 EPRNYCLARREDLDPLRHINNTINGERYGVIGRTGAKSSITLGLPRINISABEIIII-1353
DB 1303 EFRNYSTKYRENLDPLVININIVKIBCEKRYGVIGRTGAKSSTLSLALFRLIIFTEBEKIIII-1362

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:56:46 ; Search time 238.465 Seconds

(without alignments)
5594.767 Million cell updates/sec

Title: US-10-665-283-4

Perfect score: 9734

Sequence: 1 MALRGFCASDGLPMDMNV.....NTIKVPTPLCTARQLDSDRS 1891

Scoring table:

BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: Uniprot_05.80.*
2: uniprot_sprot.*
3: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7860	80.7	1531	1 MRPL_HUMAN	P33527 homo sapien
2	7769	79.8	1515	2 O9UQ9_HUMAN	O9UQ99 homo sapien
3	7695	79.1	1531	1 O86493_MACFA	O86493 macaca fasc
4	7689	79.0	1531	2 O86450_MACFA	O86450 macaca fasc
5	7444	76.5	1459	2 O9UQ97_HUMAN	O9UQ97 homo sapien
6	7419.5	76.2	1456	2 O9UQ40_HUMAN	O9UQ40 homo sapien
7	7272	74.7	1531	2 O6UR05_CANFA	O6UR05 canis famil
8	7242	74.4	1439	2 O59G19_HUMAN	O59G19 homo sapien
9	7230.5	74.3	1530	2 O8HX05_BOVIN	O8HX05 bos taurus
10	7104.5	73.0	1400	2 O9UQ98_HUMAN	O9UQ98 homo sapien
11	7002.5	71.9	1528	1 MRPL_MOUSE	O35379 mus musculu
12	6932.5	71.2	1532	2 O810E4_RAT	O810E4 rattus norv
13	6932.5	71.2	1532	2 O8CG09_RAT	O8CG09 rattus norv
14	6892	70.8	1523	2 O810G9_RAT	O810G9 rattus norv
15	6301	64.7	1525	2 O5F364_CHICK	O5F364 gallus gall
16	6075	62.4	1215	2 O68CP7_HUMAN	O68CP7 homo sapien
17	4526	46.5	1522	2 O5SUP4_MOUSE	O5SUP4 mus musculu
18	4522	46.5	1519	2 O80ZK8_MOUSE	O80ZK8 mus musculu
19	4522	46.5	1522	2 O56PH0_MOUSE	O56PH0 mus musculu
20	4508.5	46.3	1523	2 O59D10_MOUSE	O59D10 mus musculu
21	4481.5	46.1	1527	1 MRP3_HUMAN	O15438 homo sapien
22	4484.5	46.1	1533	2 O59H05_HUMAN	O59H05 homo sapien
23	4422	45.4	1514	2 O96Q49_HUMAN	O96Q49 homo sapien
24	4402	45.2	1498	2 O59DK9_MOUSE	O59DK9 mus musculu
25	4396	45.2	1522	1 MRP3_RAT	O88563 rattus norv
26	4148.5	42.6	1496	2 O4S7R7_TETNG	O4S7R7 tetradon n
27	3885.5	39.9	1548	2 O7KTC3_DROME	O7KTC3 drosophila
28	3880.5	39.9	1548	2 O7KTC1_DROME	O7KTC1 drosophila
29	3879.5	39.9	1548	2 O7KTB1_DROME	O7KTB1 drosophila
30	3874.5	39.8	1548	2 O7KTC2_DROME	O7KTC2 drosophila
31	3860.5	39.7	1548	2 O917N0_DROME	O917N0 drosophila

32	3857	39.6	1549	2 O7KTC0_DROME	O7KTC0 drosophila
33	3845.5	39.5	1548	2 O8T9C5_DROME	O8T9C5 drosophila
34	3838.5	39.4	1548	2 O7KTB9_DROME	O7KTB9 drosophila
35	3836.5	39.4	1548	2 O7KTB9_DROME	O7KTB9 drosophila
36	3818.5	39.2	1548	2 O7KTC0_DROME	O7KTC0 drosophila
37	3813.5	39.2	1548	2 O7KTC4_DROME	O7KTC4 drosophila
38	3812.5	39.2	1548	2 O7KTC4_DROME	O7KTC4 drosophila
39	3807.5	39.1	1548	2 O7KTC9_DROME	O7KTC9 drosophila
40	3793.5	39.0	1548	2 O7KTC7_DROME	O7KTC7 drosophila
41	3790	38.9	1549	2 O9VK56_DROME	O9VK56 drosophila
42	3771.5	38.7	1548	2 O7KTC5_DROME	O7KTC5 drosophila
43	3769.5	38.7	1548	2 O7KTC6_DROME	O7KTC6 drosophila
44	3741	38.4	1564	2 O8OC98_RAJER	O8OC98 raja erinac
45	3738.5	38.4	1567	2 O6PH26_BRARE	O6PH26 brachydanio

ALIGNMENTS

RESULT 1
MRPL_HUMAN STANDARD; PRT; 1531 AA.
ID MRPL_HUMAN
AC P33527; O14819; P78419;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Multidrug resistance-associated protein 1 (ATP-binding cassette sub-family C member 1).
GN Name=ABCC1; Synonyms=MRP, MRPL;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93088080; PubMed=1360704;
RA Cole S.P.C., Bhargava G., Gerlach J.H., Mackie J.E., Grant C.E., Almqvist K.C., Stewart A.J., Kurz E.U., Duncan A.M.V., Deeley R.G.;
RT "Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line.";
RL Science 258:1650-1654(1992).
RN [2]
RP SEQUENCE REVISION.
RX MEDLINE=93262415; PubMed=8098549;
RA Cole S.P.C., Deeley R.G.;
RT "Multidrug resistance-associated protein: sequence correction.";
RL Science 260:879-879(1993).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96251691; PubMed=8649356;
RA Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P.C., Deeley R.G.;
RT "Structure and expression of the messenger RNA encoding the murine multidrug resistance and expression of the messenger RNA encoding the murine multidrug resistance protein, an ATP-binding cassette transporter.";
RL Mol. Pharmacol. 49:962-971(1996).
RN [4]
RP NUCLEOTIDE SEQUENCE, AND ALTERNATIVE SPLICING.
RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;
RA Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;
RT "Analysis of the intron-exon organization of the human multidrug-resistance protein gene (MRP) and alternative splicing of its mRNA.";
RL Genomics 45:368-378(1997).
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA] OF 1131-1531.
RX MEDLINE=99455270; PubMed=10493829; DOI=10.1006/geno.1999.5927;
RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R., Fuhrmann U., Mason T., Crosby M.L., Barnstead M., Cronin L., Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S., Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
RT "Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q.";
RL Genomics 60:295-308(1999).

[6]
RT TOPOLOGY, AND CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=97442425; PubMed=9295302; DOI=10.1074/jbc.272.38.23623;
RA Helfer D.R., Altmajst K.C., Leslie E.M., Gerlach J.H., Grant C.E.,
Deeley R.G., Cole S.P.C.;
RT "Membrane topology of the multidrug resistance protein (MRP). A study
of glycosylation-site mutants reveals an extracytosolic NH2
terminus.";
RL J. Biol. Chem. 272:23623-23630(1997).
RN [7]
RN TOPOLOGY.
RX MEDLINE=97476249; PubMed=9334225; DOI=10.1074/jbc.272.42.26479;
RA Kast C., Gros P.;
RT "Topology mapping of the amino-terminal half of multidrug resistance-
RT associated protein by epitope insertion and immunofluorescence.";
RL J. Biol. Chem. 272:26479-26487(1997).
RN [8]
RN TOPOLOGY.
RX MEDLINE=98153110; PubMed=9485377; DOI=10.1021/bi972332v;
RA Kast C., Gros P.;
RT "Epitope insertion favors a six transmembrane domain model for the
carboxy-terminal portion of the multidrug resistance-associated
RT protein.";
RL Biochemistry 37:2305-2313(1998).
RN [9]
RN MUTAGENESIS OF ASP-792; ASP-793; LYS-1333 AND 1454-ASP-GLU-1455.
RX MEDLINE=21362977; PubMed=11469806; DOI=10.1006/abbi.2001.2441;
RA Cui L., Hou Y.-X., Rjordani J.R., Chang X.-B.;
RT "Mutations of the Walker B motif in the first nucleotide binding
domain of multidrug resistance protein MRP1 prevent conformational
RT maturation.";
RL Arch. Biochem. Biophys. 392:153-161(2001).
RN [10]
RN MUTAGENESIS OF TRP-1246.
RX MEDLINE=21238219; PubMed=11278867; DOI=10.1074/jbc.M011246200;
RA Ito K., Olsen S.L., Qiu W., Deeley R.G., Cole S.P.C.;
RT "Mutation of a single conserved tryptophan in multidrug resistance
RT protein 1 (MRP1/ABCC1) results in loss of drug resistance and
selective loss of organic anion transport.";
RL J. Biol. Chem. 276:15616-15624(2001).
RN [11]
RN VARIANTS GLN-633 AND VAL-671.
RX MEDLINE=20296300; PubMed=10835642; DOI=10.1038/76102;
RA Le Saux O., Urban Z., Tschuch C., Ceiszar K., Bacchelli B.,
Quaglini D., Pasquali-Bonchetti I., Pope F.M., Richards A., Terry S.,
Bercovich L., de Paeppe A., Boyd C.D.;
RT "Mutations in a gene encoding an ABC transporter cause pseudoxanthoma
RT elasticum.";
RL Nat. Genet. 25:223-227(2000).
RN [12]
RN VARIANT VAL-671.
RX MEDLINE=20283940; PubMed=10811882; DOI=10.1073/pnas.100041297;
RA Ringelstein F., Lehwahl M.G., Cristiano A.M., Uitto J.;
RT "Pseudoxanthoma elasticum: mutations in the MRP6 gene encoding a
transmembrane ATP-binding cassette (ABC) transporter.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:6001-6006(2000).
RN [13]
RN VARIANT SER-433, AND CHARACTERIZATION OF VARIANT VAL-671.
RX MEDLINE=21578494; PubMed=11721885; DOI=10.1007/s100380170017;
RA Conrad S., Kauffmann H.-M., Ito K., Deeley R.G., Cole S.P.C.,
Schrenk D.;
RT "Identification of human multidrug resistance protein 1 (MRP1)
RT mutations and characterization of a G671V substitution.";
RL J. Hum. Genet. 46:656-663(2001).
RN [14]
RN VARIANTS MET-117 AND LEU-1512.
RX MEDLINE=20579883; PubMed=11139250;
DOI=10.1002/1098-1004(2001)17:1<74::AID-HUMU14>3.0.CO;2-F;
RA Perdu J., Germain D.P.;
RT "Identification of novel polymorphisms in the pms and MRP1 (ABCC1)
RT genes at locus 16p13.1 and exclusion of both genes as responsible for
RT pseudoxanthoma elasticum.";
RL Hum. Mutat. 17:74-75(2001).

[15]
RN VARIANTS SER-43; ILE-73; GLN-723 AND GLN-1058.
RX MEDLINE=21163848; PubMed=11266082;
DOI=10.1097/00008571-200103000-00008;
RA Ito S., Ietiri I., Tanabe M., Suzuki A., Higuchi S., Otsubo K.;
RT "Polymorphism of the ABC transporter genes, MDR1, MRP1 and MRP2/cMOAT,
RT in healthy Japanese subjects.";
RL Pharmacogenetics 11:175-184(2001).
RN -1- FUNCTION: May participate directly in the active transport of
drugs into subcellular organelles or influence drug distribution
indirectly.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=8;
Comment=Additional isoforms seem to exist. Experimental
confirmation may be lacking for some isoforms;
Name=Allelons;
CC IsoId=P33527-1; Sequence=Displayed;
CC Name=Delexon-17;
CC IsoId=P33527-2; Sequence=VSP_000037;
CC Name=Delexon-18;
CC IsoId=P33527-3; Sequence=VSP_000038;
CC Name=Delexon-30;
CC IsoId=P33527-4; Sequence=VSP_000039;
CC Name=Delexon-17-18;
CC IsoId=P33527-5; Sequence=VSP_000037, VSP_000038;
CC Name=Delexon-17-30;
CC IsoId=P33527-6; Sequence=VSP_000037, VSP_000039;
CC Name=Delexon-18-30;
CC IsoId=P33527-7; Sequence=VSP_000038, VSP_000039;
CC Name=Delexon-17-18-30;
CC IsoId=P33527-8; Sequence=VSP_000037, VSP_000039;
CC -1- TISSUE SPECIFICITY: Lung, testis and peripheral blood mononuclear
cells.
CC -1- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.
CC -1- SIMILARITY: Contains 2 ABC transmembrane type-1 domains.
CC -1- SIMILARITY: Contains 2 ABC transporter domains.
CC -1- DATABASE: NMBE=NCIas Genet. Cytogenet. Oncol. Haematol.;
WWW="http://www.inbiochem.fr/services/chromocancer/genes/MRP106.html".
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
DR EMBL, L05628; AAB4616.1; -, mRNA.
DR EMBL, AF022853; AAB83983.1; -, Genomic DNA.
DR EMBL, AF022824; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022825; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022826; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022827; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022828; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022829; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022830; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022831; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022832; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022833; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022834; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022835; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022836; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022837; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022838; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022839; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022840; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022841; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022842; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022843; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022844; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022845; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022846; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022847; AAB83983.1; JOINED; Genomic DNA.
DR EMBL, AF022848; AAB83983.1; JOINED; Genomic DNA.

DR EMBL: AF022849; AAB83983.1; JOINED; Genomic DNA.
DR EMBL: AF022850; AAB83983.1; JOINED; Genomic DNA.

Query Match 80.7%; Score 7860; DB 1; Length 1531;

Best Local Similarity 100.0%; Pident. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MALRGFCSADGSDPLMDMNTVNTSNDFTCFQNTLVWVPCFYLMACPFYLYSRH 60
DB 1 MALRGFCSADGSDPLMDMNTVNTSNDFTCFQNTLVWVPCFYLMACPFYLYSRH 60
QY 61 DRGVIQMTPLNKTALGFLIMIVCMADLFYSFMEWRSRGIFLAFLVPSPTLIGITLLA 120
DB 61 DRGVIQMTPLNKTALGFLIMIVCMADLFYSFMEWRSRGIFLAFLVPSPTLIGITLLA 120
QY 121 TFLIOLERRKGVSSGIMLTFWALVALCALILRSKIMTALKEDAQVDLFFDITFYVYS 180
DB 121 TFLIOLERRKGVSSGIMLTFWALVALCALILRSKIMTALKEDAQVDLFFDITFYVYS 180
QY 181 LLLIQLVLSGFSDBSPFSETHDNPCESSASFLSRTIFPMWTGLIVRGYRPLEGSD 240
DB 181 LLLIQLVLSGFSDBSPFSETHDNPCESSASFLSRTIFPMWTGLIVRGYRPLEGSD 240
QY 241 LMSLNKEDTSEOVVPLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKVDANEYEAL 300
DB 241 LMSLNKEDTSEOVVPLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKVDANEYEAL 300
QY 301 IVKSPQKEMNPSLFKVLKTFGRPYFLMSFFPKAIDLMFSGPOLIKLIKPVNDTAPD 360
DB 301 IVKSPQKEMNPSLFKVLKTFGRPYFLMSFFPKAIDLMFSGPOLIKLIKPVNDTAPD 360
QY 361 WQGFYTVLVFVACLOTLVHOVPHICFSGMRKIKTAVIGAVYRKALVTNSRKSTV 420
DB 361 WQGFYTVLVFVACLOTLVHOVPHICFSGMRKIKTAVIGAVYRKALVTNSRKSTV 420
QY 421 GEIVNLMSVDAQRFMDATYINMISAPLOVITLALYLLMNLGSPVLAVVWLVMPVN 480
DB 421 GEIVNLMSVDAQRFMDATYINMISAPLOVITLALYLLMNLGSPVLAVVWLVMPVN 480
QY 481 AVNANKTKTYQVAHMSKDNRIKLMNEILINGIKVLKYLAMELAFKDVLAIRQELKVLK 540
DB 481 AVNANKTKTYQVAHMSKDNRIKLMNEILINGIKVLKYLAMELAFKDVLAIRQELKVLK 540
QY 541 KSAVLAVGTFVWCPELVALCTFAVYVITDENIIDAQTAFTSLAPNITRPLNLP 600
DB 541 KSAVLAVGTFVWCPELVALCTFAVYVITDENIIDAQTAFTSLAPNITRPLNLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHELEBDSIERRPVKDGGGNSITVYVATFTWARSDEPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHELEBDSIERRPVKDGGGNSITVYVATFTWARSDEPT 660
QY 661 LNSITFSPREGALVAVVGVCCKGSSLSALLAEMDVKEGHVAKGSVAVYVPOQAWIQND 720
DB 661 LNSITFSPREGALVAVVGVCCKGSSLSALLAEMDVKEGHVAKGSVAVYVPOQAWIQND 720
QY 721 SLAENILFGQLEBPYRSYIOACALLPDLIELPSGRTEIGEGVNLSSGQKORVSLAR 780
DB 721 SLAENILFGQLEBPYRSYIOACALLPDLIELPSGRTEIGEGVNLSSGQKORVSLAR 780
QY 781 AAVSNADITYLPDPLSAVDHVGKHPENYIGPKMLKNTIRILVTHSMGYLPQVDYIV 840
DB 781 AAVSNADITYLPDPLSAVDHVGKHPENYIGPKMLKNTIRILVTHSMGYLPQVDYIV 840
QY 841 MSGGKISEMGSYQELLARDGAFAPFLRTYASTEOBQDAEENGVTGVSGPKKAKOMENGM 900
DB 841 MSGGKISEMGSYQELLARDGAFAPFLRTYASTEOBQDAEENGVTGVSGPKKAKOMENGM 900
QY 901 LVYDSAGKOLOROLSSSSSGDISRHHNSTAEIQKAEKEETWKLMEADKQGTQGVKL 960
DB 901 LVYDSAGKOLOROLSSSSSGDISRHHNSTAEIQKAEKEETWKLMEADKQGTQGVKL 960
QY 961 SYVWDYKAIQGLFISFLIFMCONHVSALASNWYLSLMTDDPIVNGTOEHTKVRLSVYG 1020
DB 961 SYVWDYKAIQGLFISFLIFMCONHVSALASNWYLSLMTDDPIVNGTOEHTKVRLSVYG 1020
```

```
DB 961 SYVWDYKAIQGLFISFLIFMCONHVSALASNWYLSLMTDDPIVNGTOEHTKVRLSVYG 1020
QY 1021 ALGISQIAVFGYSMAVSTIGIILASRCHADLHSLTIRSPMSFEPRPSGNLVNRSKEL 1080
DB 1021 ALGISQIAVFGYSMAVSTIGIILASRCHADLHSLTIRSPMSFEPRPSGNLVNRSKEL 1080
QY 1081 DTVDMSIMEVYIKMFMSLFAVIGACIYIILATPAAIIPPLIGLYFFVORFYVASSROL 1140
DB 1081 DTVDMSIMEVYIKMFMSLFAVIGACIYIILATPAAIIPPLIGLYFFVORFYVASSROL 1140
QY 1141 KRLBSVRSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENQKAYYPSIVANWLA 1200
DB 1141 KRLBSVRSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENQKAYYPSIVANWLA 1200
QY 1201 VRLCEVNCIVLPAALFAVYSRHSLSAGLVGSYSLOVTTYIYVWVYRMSSEMETNIVA 1260
DB 1201 VRLCEVNCIVLPAALFAVYSRHSLSAGLVGSYSLOVTTYIYVWVYRMSSEMETNIVA 1260
QY 1261 VERLKEYSETKEKAPMOIOETAPSSWPQYGRVFRYVCPRYREDLDFVLRIHVTINGG 1320
DB 1261 VERLKEYSETKEKAPMOIOETAPSSWPQYGRVFRYVCPRYREDLDFVLRIHVTINGG 1320
QY 1321 EKVGIQRTGAGKSSLLTGLFRINESABGEIILIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIQRTGAGKSSLLTGLFRINESABGEIILIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGSLRNMLDPPSOYSDEEWTSLLEAHLKQFVSALPKLDHECAEGEENISVQORQVCL 1440
DB 1381 SGSLRNMLDPPSOYSDEEWTSLLEAHLKQFVSALPKLDHECAEGEENISVQORQVCL 1440
QY 1441 ARALLRKTILVLEDEATAADLETDLOSTIRTOFEDCVLTIAHNLNTIMOTRYIVL 1500
DB 1441 ARALLRKTILVLEDEATAADLETDLOSTIRTOFEDCVLTIAHNLNTIMOTRYIVL 1500
QY 1501 DKGEIOEYGAAPDLOQRGLFYSMAXDAGLV 1531
DB 1501 DKGEIOEYGAAPDLOQRGLFYSMAXDAGLV 1531

RESULT 2
Q9U099 HUMAN
ID Q9U099 HUMAN PRELIMINARY; PRT; 1515 AA.
AC Q9U099.
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Multidrug resistance protein (Fragment).
GN Name=MRP;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;
RA Grant C.E., Kurz E.U., Cole S.P.C., Deley R.G.;
RT "Analysis of the intron-exon organization of the human multidrug-
resistance protein gene (MRP) and alternative splicing of its mRNA.";
RL Genomics 45:368-378 (1997).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL: AF022853; AAB83979.1; -; Genomic DNA.
DR EMBL: AF022827; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022828; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022829; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022831; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022833; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022835; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022837; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022839; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022841; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022849; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022850; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022849; AAB83979.1; JOINED; Genomic DNA.
```

DR EMBL: AF022848; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022847; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022846; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022845; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022844; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022843; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022842; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022841; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022840; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022839; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022838; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022836; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022834; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022832; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022826; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022825; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022824; AAB83979.1; JOINED; Genomic DNA.
DR EMBL: AF022830; AAB83979.1; JOINED; Genomic DNA.
DR HSSP: P08716; IMT0.
DR Ensembl: ENSG00000103222; Homo sapiens.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0016887; F:ATPase activity; IEA.
DR GO: GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:0000166; F:nucleotide binding; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR003593; AAA ATPase.
DR InterPro: IPR011527; ABC membrane 1.
DR InterPro: IPR001140; ABC_TM_transp.
DR InterPro: IPR003439; ABC_transp_like.
DR InterPro: IPR005292; MRP assoc.
DR Pfam: PF00664; ABC_tran; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR Prodom: PD000006; ABC_transporter; 2.
DR SMART: SM00382; AAA; 2.
DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
DR PROSITE: PS00929; ABC_TMIF; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE: PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR ATP-binding; Nucleotide-binding.
DR KATP-binding; Nucleotide-binding.
DR NON_TIR 1 1
DR SEQUENCE 1515 AA; 169853 MW; DC85592817C439FE CRC64;
Query March 79.8%; Score 7769; DB 2; Length 1515;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 DMNTVMTSNPDEFTKCFQNTVLVWPCFYLMACFPFFLYLSRHDGTYIQMTPLNKTETA 76
DB 1 DMNTVMTSNPDEFTKCFQNTVLVWPCFYLMACFPFFLYLSRHDGTYIQMTPLNKTETA 60
QY 77 LGPLMTIVCWADLFYFSEWERSRGIFLAFVLVSPPTLIGITTLATFLLIQLERRKGVSSG 136
DB 61 LGPLMTIVCWADLFYFSEWERSRGIFLAFVLVSPPTLIGITTLATFLLIQLERRKGVSSG 120
QY 137 IMTLFWVAIVCALAIIIRSKTMTALKEDAOVDLPDITFFVYFSLILLQLVLSGCSDDSP 196
DB 121 IMTLFWVAIVCALAIIIRSKTMTALKEDAOVDLPDITFFVYFSLILLQLVLSGCSDDSP 180
QY 197 LFEETIHDPNCPRESSASFLSRITFWMITGLIVRGYROPLESGSDLSLNKEDTSBOVVPV 256
DB 181 LFEETIHDPNCPRESSASFLSRITFWMITGLIVRGYROPLESGSDLSLNKEDTSBOVVPV 240
QY 257 LVKWKKECAKTRKQPVYVSSKDPAPQKSSKVDANEVEVALIVKSPQKEMNSLPKVV 316
DB 241 LVKWKKECAKTRKQPVYVSSKDPAPQKSSKVDANEVEVALIVKSPQKEMNSLPKVV 300
QY 317 LYKTFGYPFLMSFFPKAIIHDLMMPSGQIILKLLIKFVNDTKAPDMOGIFTYVLLFVTAQL 376
DB 301 LYKTFGYPFLMSFFPKAIIHDLMMPSGQIILKLLIKFVNDTKAPDMOGIFTYVLLFVTAQL 360

QY 377 QTVLHGYFHI CFVSGMRITTAIVIGAVYRKALVYITNSARKSVYGEIVNLMSVDAGRFMD 436
DB 361 QTVLHGYFHI CFVSGMRITTAIVIGAVYRKALVYITNSARKSVYGEIVNLMSVDAGRFMD 420
QY 437 LATYINNIWASPIQVILATVLLMLNIGPSVLAQVAVVUMLPVNAVNAAMTKTYQVAMHK 496
DB 421 LATYINNIWASPIQVILATVLLMLNIGPSVLAQVAVVUMLPVNAVNAAMTKTYQVAMHK 480
QY 497 SKDNRIKLAMEIINGIKVLKLYAMELAFKDKVLAIROBEIKVLKKSAYLSAVGTFTWVCT 556
DB 481 SKDNRIKLAMEIINGIKVLKLYAMELAFKDKVLAIROBEIKVLKKSAYLSAVGTFTWVCT 540
QY 557 PFLVALCTFAVYVYITDENNILLDAQTAFAVSIALFNILRFPPLNIPMTVSSIVQASVSLKRL 616
DB 541 PFLVALCTFAVYVYITDENNILLDAQTAFAVSIALFNILRFPPLNIPMTVSSIVQASVSLKRL 600
QY 617 RIFLSHEELRPPDSTERPVNDGGTNSITVRNATFTWABSDPPTLNITTSIPSGALVAV 676
DB 601 RIFLSHEELRPPDSTERPVNDGGTNSITVRNATFTWABSDPPTLNITTSIPSGALVAV 660
QY 677 VGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAVVPQAMIQNDSLRENTLFGQLEEPY 736
DB 661 VGVGCGKSSLLSALLAEMDKVEGHVAIKGSVAVVPQAMIQNDSLRENTLFGQLEEPY 720
QY 737 YRSVIAQCALLPDLEILPSSGDRTEIGKGVNLSGGQKQSVSLARAVYSNADIVLPDPLS 796
DB 721 YRSVIAQCALLPDLEILPSSGDRTEIGKGVNLSGGQKQSVSLARAVYSNADIVLPDPLS 780
QY 797 AYDAHVGRKHIFENVIGRKGMLKXKTRILYTHMSYLPQVUVITYMSGGKISEMGSYQEL 856
DB 781 AYDAHVGRKHIFENVIGRKGMLKXKTRILYTHMSYLPQVUVITYMSGGKISEMGSYQEL 840
QY 857 ARDGAFAEFLRTYAATQEOQDAEENGTVGSGPEKAKQKQENGLVYDSAGKQLOROLSS 916
DB 841 ARDGAFAEFLRTYAATQEOQDAEENGTVGSGPEKAKQKQENGLVYDSAGKQLOROLSS 900
QY 917 SSSYSGDISRRHNSTAELOAKAEKKEBTWKLMEADKAQTCQVKLSYVWYDKAIGLFISF 976
DB 901 SSSYSGDISRRHNSTAELOAKAEKKEBTWKLMEADKAQTCQVKLSYVWYDKAIGLFISF 960
QY 977 LSTFLPNCNVSAIASNYWMLSLWTDPIVNGTOGHTVRLSYVCAIGISQIAIFGVSMA 1036
DB 961 LSTFLPNCNVSAIASNYWMLSLWTDPIVNGTOGHTVRLSYVCAIGISQIAIFGVSMA 1020
QY 1037 VSIIGILASRCLAHVDLHSLIRSPMSFFERTPSGNLVNRSKELDYDMSIPEVIKAFMG 1096
DB 1021 VSIIGILASRCLAHVDLHSLIRSPMSFFERTPSGNLVNRSKELDYDMSIPEVIKAFMG 1080
QY 1097 SLEFNVIGACIYIILLATPIAIIIPPLGLIYFVQRFYVASSRQIKRLSEVSRSFVSHFN 1156
DB 1081 SLEFNVIGACIYIILLATPIAIIIPPLGLIYFVQRFYVASSRQIKRLSEVSRSFVSHFN 1140
QY 1157 ETLIGSVIYAFEEQERFIHOSDLKVDENQKAYVPSIVANRMLAVRLECVCNCIVLPAAL 1216
DB 1141 ETLIGSVIYAFEEQERFIHOSDLKVDENQKAYVPSIVANRMLAVRLECVCNCIVLPAAL 1200
QY 1217 PAVISRSLSAGVLGVSYSLOVTTYINMLVRMSSEMETNIYAVERLKESESETEKAPW 1276
DB 1201 PAVISRSLSAGVLGVSYSLOVTTYINMLVRMSSEMETNIYAVERLKESESETEKAPW 1260
QY 1277 QIQETAPSSWPQVGRVFRNYCLRYREDLDFVLRIHNVITNGEKVGI VGRTAGKSSL 1336
DB 1261 QIQETAPSSWPQVGRVFRNYCLRYREDLDFVLRIHNVITNGEKVGI VGRTAGKSSL 1320
QY 1337 TLGLPRINESAGEIITIDGINIAKIGLHDLRFKTTIIPDPVLFSGSIRNMLDPFSQYSD 1386
DB 1321 TLGLPRINESAGEIITIDGINIAKIGLHDLRFKTTIIPDPVLFSGSIRNMLDPFSQYSD 1380
QY 1397 EEWVTSLELAHLKQFVSALPDKLDHECABEGGENISVGOBOLVCLARALLRXTKTLIVDEA 1456
DB 1381 EEWVTSLELAHLKQFVSALPDKLDHECABEGGENISVGOBOLVCLARALLRXTKTLIVDEA 1440
QY 1457 TAAVDLETDLLIOSTIRTOFEDCTVLITAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQ 1516

Db 1441 TAAVDELTDLIQSTIRQFEDCTVLTIAHRLNTIMDYTRVYLDKGEIQGYGAPSDLLQ 1500
 Qy 1517 ORGLFYSMADAGLV 1531
 Db 1501 ORGLFYSMADAGLV 1515

RESULT 3

0864R9_MACFA PRELIMINARY; PRT; 1531 AA.
 ID 0864R9_MACFA
 AC 0864R9;
 DT 01-JUN-2003 (TREMBlurel. 24, Created)
 DT 01-JUN-2003 (TREMBlurel. 24, Last sequence update)
 DT 01-MAR-2004 (TREMBlurel. 26, Last annotation update)
 DE Multidrug resistance protein 1B.
 GN Name=MRP1;
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 OC Cercopithecoidea; Cercopithecinae; Macaca.
 OC NCBI_TaxID=9541;
 RX MEDLINE=22544876; PubMed=12657726;
 RA Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,
 RA Dantzig A.H., Perry W.L.;
 RT "Cloning and functional characterization of the multidrug resistance-
 RT associated protein (MRP1/ABCC1) from the cynomolgus monkey.";
 RL Mol. Cancer Ther. 2:307-316(2003).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.

RA Perry W.L., III, Godinot N.; EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the ABC transporter family.
 DR EMBL; AY146673; AAN65349.1; -; mRNA.
 DR HSSP; P08716; IMTO.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0016887; F:ATPase activity; IEA.
 DR GO; GO:0042656; F:ATPase activity, coupled to transmembrane m. . .; IEA.
 DR GO; GO:0000166; F:nucleotide binding; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR001593; AAA_ATPase.
 DR InterPro; IPR011527; ABC_membrane_1.
 DR InterPro; IPR001140; ABC_TM_transp.
 DR InterPro; IPR003439; ABC_transp_like.
 DR InterPro; IPR005292; MRP_assoc.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00664; ABC_membrane; 2.
 DR Pfam; PF00005; ABC_tran; 2.
 DR ProDom; PD000006; ABC_transporter; 2.
 DR SMART; SM00382; AAA; 2.
 DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
 DR PROSITE; PS50929; ABC_TMIF; 2.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
 DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 KW ATP-binding; Nucleotide-binding; Repeat.
 SQ SEQUENCE 1531 AA; 171659 MW; 1AE78BEFDF9EF459 CRC64;

Query Match 79.1%; Score 7695; DB 2; Length 1531;
 Best Local Similarity 97.5%; Pred. No. 0;
 Matches 1493; Conservative 20; Mismatches 18; Indels 0; Gaps 0;

Qy 1 MALRGFCADGSDPLMDMNTVNTSNDPTKCFONTVLVWPCFYLMACEPFYFLYSRH 60
 Db 1 MALRGFCADGSDPLMDMNTVNTSNDPTKCFONTVLVWPCFYLMACEPFYFLYSRH 60
 Qy 61 DRGVIQMTPLNKTALGFLIMVYCADLFYSFWEBSRGIFLAVPLVSPTLGITTLLA 120
 Db 61 DRGVIQMTPLNKTALGFLIMVYCADLFYSFWEBSRGIFLAVPLVSPTLGITTLLA 120

Qy 121 TFLIQERRKGVSSGIMLTFWLVAVCALAILRSKIMTALKEDAOVDFRDTFFYYFS 180
 Db 121 TFLIQERRKGVSSGIMLTFWLVAVCALAILRSKIMTALKEDAOVDFRDTFFYYFS 180
 Qy 181 LLLIQVLVSCFSDRSPLESETIHDNPNCPESSASFLSRITFWWTGLIVRGYRPLEGSD 240
 Db 181 LVLIIQVLVSCFSDRSPLESETIHDNPNCPESSASFLSRITFWWTGLIVRGYRPLEGSD 240
 Qy 241 LMSLNKEDTSEOVVPUVVKMKKECATRQKQPVVNVVSSVDPAPKSSSKVDANEVEAL 300
 Db 241 LMSLNKEDTSEOVVPUVVKMKKECATRQKQPVVNVVSSVDPAPKSSSKVDANEVEAL 300
 Qy 301 IVKSPQKEMPSPLEKVLVYKTFGPFYFLMSFFFKAIHDMFMFSGPOILLKILKFVNDTAPD 360
 Db 301 IVKSPQKEMPSPLEKVLVYKTFGPFYFLMSFFFKAIHDMFMFSGPOILLKILKFVNDTAPD 360
 Qy 361 WQGYFYTALLFVTAQLOTLVLAHQYFHIQFVSGMRIKTAVI GA VYRKALVITNSARKSTV 420
 Db 361 WQGYFYTALLFVTAQLOTLVLAHQYFHIQFVSGMRIKTAVI GA VYRKALVITNSARKSTV 420
 Qy 421 GEIVNLMVSVDAPRPMDLATTYNNMWSAPLOYIALLYLMLNLGFSVLGA VAVMLAMPVN 480
 Db 421 GEIVNLMVSVDAPRPMDLATTYNNMWSAPLOYIALLYLMLNLGFSVLGA VAVMLAMPVN 480
 Qy 481 AVNAMKTKTYQVAHMKSKDKNRILKLMNLIIGIKVLKLYAREIAFKDKVLAIROBELKYLK 540
 Db 481 AVNAMKTKTYQVAHMKSKDKNRILKLMNLIIGIKVLKLYAREIAFKDKVLAIROBELKYLK 540
 Qy 541 KSAVLASAVGTFTWCTPPLVALCTFAVYVTTIDENNILDAOTAFAVSLALFNILRPLNLP 600
 Db 541 KSAVLASAVGTFTWCTPPLVALCTFAVYVTTIDENNILDAOTAFAVSLALFNILRPLNLP 600
 Qy 601 MTISSIVQASVSLKRLRILFISHBELPDSIERRPVQDGGTNSITVNAATFTMARSDPT 660
 Db 601 MTISSIVQASVSLKRLRILFISHBELPDSIERRPVQDGGTNSITVNAATFTMARSDPT 660
 Qy 661 LNTGTFSPRGALVAVVGVGCGKSLSLALBMDVBEHVAKSGVA VVPOQAMQND 720
 Db 661 LNTGTFSPRGALVAVVGVGCGKSLSLALBMDVBEHVAKSGVA VVPOQAMQND 720
 Qy 721 SLRENILFGQLEBPYRSYIOACALPDEILIPSGDRTIEGKVMASGQOKORVSLAR 780
 Db 721 SLRENILFGQLEBPYRSYIOACALPDEILIPSGDRTIEGKVMASGQOKORVSLAR 780
 Qy 781 AVYSNADTYFDPLSAVDAHVGHIFENYIGPKMKKNTRIILVTHSMGYLPQVDVIIV 840
 Db 781 AVYSNADTYFDPLSAVDAHVGHIFENYIGPKMKKNTRIILVTHSMGYLPQVDVIIV 840
 Qy 841 MSGGKISEMGSYOELLARDGAPAEPLRTVASTOEOAEENGVTGVS GPKKAKOMENG 900
 Db 841 MSGGKISEMGSYOELLARDGAPAEPLRTVASTOEOAEENGVTGVS GPKKAKOMENG 900
 Qy 901 LVYDSAGKQLOROLSSSSYSYSDISRHHNSTAELOKAEAKKEETWKLMEADKAQTQGVKL 960
 Db 901 LVYDSAGKQLOROLSSSSYSYSDISRHHNSTAELOKAEAKKEETWKLMEADKAQTQGVKL 960
 Qy 961 SVYWDYKAIAGLIFSLIFSLIFMCHNVSAIASNYWLSLMTDPIVNGTQHTYKRLSVYG 1020
 Db 961 SVYWDYKAIAGLIFSLIFSLIFMCHNVSAIASNYWLSLMTDPIVNGTQHTYKRLSVYG 1020
 Qy 1021 ALGISGCIANFGVSMNVSTIGLILASRCLAHVDLHSLIRSPMSFFERPSGNLVNRPSEKL 1080
 Db 1021 ALGISGCIANFGVSMNVSTIGLILASRCLAHVDLHSLIRSPMSFFERPSGNLVNRPSEKL 1080
 Qy 1081 DTVDSMIPFVYKMFMSGLFNVIGACIVYLLATPILAIILPPLGLIYFFVORFYAASSROL 1140
 Db 1081 DTVDSMIPFVYKMFMSGLFNVIGACIVYLLATPILAIILPPLGLIYFFVORFYAASSROL 1140
 Qy 1141 KRLESYSRSPVYSHFNMTLLGVSVITRAFEEQERFIHOSDLKVDENOKAYYPSIVANRWLA 1200
 Db 1141 KRLESYSRSPVYSHFNMTLLGVSVITRAFEEQERFIHOSDLKVDENOKAYYPSIVANRWLA 1200
 Qy 1201 VRLCEVCNIVLPAALFAVISRSLSGVLGVSYSLSQVTTYTLNMLVRRMSEMETNIVA 1260

Db 1201 VRLECVNCLVLFALFAVLSRHSLSAGLVGLSVSYSLQVTTYTNMLVRMSSEMETIVA 1260
Qy 1261 VERLKESETEKEAPMOIQETAPSSWMPQVGRVEFRNCLRYREDLDFVLHINVTNGG 1320
Db 1261 VERLKESETEKEAPMOIQETAPSSWMPQVGRVEFRNCLRYREDLDFVLHINVTNGG 1320
Qy 1321 EKVGIIVRTGAGKSLTLGLFRINESAGEIIGDINIAKIGLHDLRFKTIIPQDPVLF 1380
Db 1321 EKVGIIVRTGAGKSLTLGLFRINESAGEIIGDINIAKIGLHDLRFKTIIPQDPVLF 1380
Qy 1381 SSGIRNMLDPPSQSDEEVTWLSLELAHKDFVSALPDLHDECAAGGNLSVGQOLVCL 1440
Db 1381 SSGIRNMLDPPSQSDEEVTWLSLELAHKDFVSALPDLHDECAAGGNLSVGQOLVCL 1440
Qy 1441 ARALLRKTIIVLEATAVDLJETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVVL 1500
Db 1441 ARALLRKTIIVLEATAVDLJETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVVL 1500
Qy 1501 DKGEIOEYGAPSDLLQGRGLFYMAKAGLV 1531
Db 1501 DKGEIOEYGAPSDLLQGRGLFYMAKAGLV 1531

RESULT 4

Q084S0_MACFA PRELIMINARY; PRT; 1531 AA.
ID Q084S0;
AC Q084S0;
DT 01-JUN-2003 (TREMblrel. 24, Created)
DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE Multidrug resistance protein 1A.
GN Name=MRP1;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecidae; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RA MEDLINE=22544876; PubMed=12657726;
RA Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,
RA Dantzig A.H., Perry W.L.;
RT "Cloning and functional characterization of the multidrug resistance-
RT associated protein (MRP1/ABCC1) from the cynomolgus monkey.";
RL Mol. Cancer Ther. 2:307-316(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RA Perry W.L., Ili, Godinot N.;
RL Submitted (AUG-2002) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY146672; AAN5348.1; -; mRNA.
DR HSSP; P08716; IMT0.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS09299; ABC_TMIF; 2.

DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS05093; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS0107; PROTEIN_KINASE_ATP; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1531 AA; 171701 MW; 16871286CC2DB89 CRC64;

Query Match 79.0%; Score 7689; DB 2; Length 1531;
Best Local Similarity 97.6%; Pred. No. 0;
Matches 1494; Conservative 18; Mismatches 19; Indels 0; Gaps 0;

Qy 1 MALRGFCSADGSDPLMDMNTWTNTNSNDFPKCFQNTLVWVPCGYLWACPFYFLYSRH 60
Db 1 MALRGFCSADGSDPLMDMNTWTNTNSNDFPKCFQNTLVWVPCGYLWACPFYFLYSRH 60
Qy 61 DRGIQMTPLNKTKTALGFLIMIVCMADLFYSFMRSGIFLAVPLVSPITLITLLA 120
Db 61 DRGIQMTPLNKTKTALGFLIMIVCMADLFYSFMRSGIFLAVPLVSPITLITLLA 120
Qy 121 TFLIQERRKGVSSGIMLTFWLVAVCALILRSKIMTALKEDAVQVDFPDITFYFYS 180
Db 121 TFLIQERRKGVSSGIMLTFWLVAVCALILRSKIMTALKEDAVQVDFPDITFYFYS 180
Qy 181 LLLIQVLSCFSDSPLESETIHDNPPCPSSASFLSRTFMWITGLIVRGYRQPLSGSD 240
Db 181 LLLIQVLSCFSDSPLESETIHDNPPCPSSASFLSRTFMWITGLIVRGYRQPLSGSD 240
Qy 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPAPKSSKVDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVVYSSKDPAPKSSKVDANEVEAL 300
Qy 301 IVKSPQKBNBSLFKVIYKTFGPYFLMSFFPKAIHDLMMSSGPQILKILKFVNDYKAPD 360
Db 301 IVKSPQKBNBSLFKVIYKTFGPYFLMSFFPKAIHDLMMSSGPQILKILKFVNDYKAPD 360
Qy 361 WQGYFYVLLVFYTKCLCTIVLHOYFHI CFSGMIKTAIVGAVYRKALVITNSARKSTV 420
Db 361 WQGYFYVLLVFYTKCLCTIVLHOYFHI CFSGMIKTAIVGAVYRKALVITNSARKSTV 420
Qy 421 GEIVNLSVDAQREMDLATYINMISAPLQVITLALYLLMLAGPSVLAVGAVMYLVNPN 480
Db 421 GEIVNLSVDAQREMDLATYINMISAPLQVITLALYLLMLAGPSVLAVGAVMYLVNPN 480
Qy 481 AVMAKTKTYOVAMHKSNDNRKILMNEILNGIKVLKYAMELAFKDYALAIROBELKYLK 540
Db 481 AVMAKTKTYOVAMHKSNDNRKILMNEILNGIKVLKYAMELAFKDYALAIROBELKYLK 540
Qy 541 KSAVLSAVGFTTWCCTPEFVALCTFAVYVTDENNIIIDAOAFVSLAFNLRPNTLP 600
Db 541 KSAVLSAVGFTTWCCTPEFVALCTFAVYVTDENNIIIDAOAFVSLAFNLRPNTLP 600
Qy 601 MVISIVQASVSLKRLNIFLSHELEPDSIERRPVKQGGTNSITVRNATPTMARSDPT 660
Db 601 MVISIVQASVSLKRLNIFLSHELEPDSIERRPVKQGGTNSITVRNATPTMARSDPT 660
Qy 661 LNGTHPSIPBGALVAVVGVQVCGKSSLLSALLAEMDVKEGHVAKGSVAVVYVQQAIIQND 720
Db 661 LNGTHPSIPBGALVAVVGVQVCGKSSLLSALLAEMDVKEGHVAKGSVAVVYVQQAIIQND 720
Qy 721 SLRENILFGCLLEPPYRSVYQAALPLDLEILSGBRTEIGEGVNLSSGQKRVSLAR 780
Db 721 SLRENILFGCLLEPPYRSVYQAALPLDLEILSGBRTEIGEGVNLSSGQKRVSLAR 780
Qy 781 AVYENADIVLEPDDPLSAVDHVGKHIFENVIGPGMLKNKRIIVYTHSMSTYLPQVDVIV 840
Db 781 AVYENADIVLEPDDPLSAVDHVGKHIFENVIGPGMLKNKRIIVYTHSMSTYLPQVDVIV 840
Qy 841 MSGKISMSGSYOELIARDGAFALFRTYASTEQEODAEENGVTGVSGPGKEAKOMENG 900
Db 841 MSGKISMSGSYOELIARDGAFALFRTYASTEQEODAEENGVTGVSGPGKEAKOMENG 900
Qy 901 LVYDSAGKOLOROLSSSSSYSGDLSRHNSYTAELQKAKKEETWKLMEADKATQGVYKL 960
Db 901 LVYDSAGKOLOROLSSSSSYSGDLSRHNSYTAELQKAKKEETWKLMEADKATQGVYKL 960

Db 301 LYKTEGFEYFLMSFEFFKAIDHLMFSGPOLIKLLIKFVNDTKAPDMQGYFTVILLFVTAOL 360
Qy 377 QTLVHOYFHICFSGMGRKIKTAVIGAVYRKALVITNSARKSSTVGEIYNLMSVDAQRFMD 436
Db 361 QTLVHOYFHICFSGMGRKIKTAVIGAVYRKALVITNSARKSSTVGEIYNLMSVDAQRFMD 420
Qy 437 LATYNNIWSAPLOVILLALYLLMLNGPSVLAVAVMLAVPVNAVAMAKTKTYQVAHMK 496
Db 421 LATYNNIWSAPLOVILLALYLLMLNGPSVLAVAVMLAVPVNAVAMAKTKTYQVAHMK 480
Qy 497 SKONRIKLAMEILLNGIKVLKLYAMELAPKDKVLAIRQELKVLKKSATLSAVGFTWVCT 556
Db 481 SKONRIKLAMEILLNGIKVLKLYAMELAPKDKVLAIRQELKVLKKSATLSAVGFTWVCT 540
Qy 557 PFLVALCTFAVYVITIDENNIIIDAOTAFVSLFNLIRPPLNILPMVSISSIVQASVSLKRL 616
Db 541 PFLVALCTFAVYVITIDENNIIIDAOTAFVSLFNLIRPPLNILPMVSISSIVQASVSLKRL 600
Qy 617 RIFLSHEELEBDSIERPPVKDGGTNSITVRNATFTWARSDEPTLNGITFSIPGALVAV 676
Db 601 RIFLSHEELEBDSIERPPVKDGGTNSITVRNATFTWARSDEPTLNGITFSIPGALVAV 660
Qy 677 VGVQCGCKSSILSLALLAEMDKVEGHVAIKGSVAVVPQAWIQNDSIRENIIIFGQLEBPY 736
Db 661 VGVQCGCKSSILSLALLAEMDKVEGHVAIKGSVAVVPQAWIQNDSIRENIIIFGQLEBPY 720
Qy 737 YRSYIQAACALPDLEIILPSGDRTEIGEXGVMLSGGQKQVSLARAVYSNADIYLFDDPLS 796
Db 721 YRSYIQAACALPDLEIILPSGDRTEIGEXGVMLSGGQKQVSLARAVYSNADIYLFDDPLS 747
Qy 797 AVDAHVGKHLFENVIGPKMKLNKTRILIVTHSMSYLPQVDVIIVMSGKISEMGSYOELL 856
Db 748 -----KTRILIVTHSMSYLPQVDVIIVMSGKISEMGSYOELL 784
Qy 857 ARDAPAFELRTYASTBEOBDAFENGVTGSGPGKEAQMNGMLYTDSACKQIQRQLSS 916
Db 785 ARDAPAFELRTYASTBEOBDAFENGVTGSGPGKEAQMNGMLYTDSACKQIQRQLSS 844
Qy 917 SSSYSGDISSRHNSHTAELOKAEAKKEETWKLEADKAOTGOVKLSVYDWYKAIQGLFISF 976
Db 845 SSSYSGDISSRHNSHTAELOKAEAKKEETWKLEADKAOTGOVKLSVYDWYKAIQGLFISF 904
Qy 977 LSIPLFCMNHVSALASNYWLSLMTDDPIVNGTOEHTKVRLSVYGAIGISQGIAVFGYSMA 1036
Db 905 LSIPLFCMNHVSALASNYWLSLMTDDPIVNGTOEHTKVRLSVYGAIGISQGIAVFGYSMA 964
Qy 1037 VSIQGIILASRCLAHVDLHSHIRSPWSFPERTPSGNLVNRFSKELDTVDSMIPEVILKXEMG 1096
Db 965 VSIQGIILASRCLAHVDLHSHIRSPWSFPERTPSGNLVNRFSKELDTVDSMIPEVILKXEMG 1024
Qy 1097 SLFVNTGACIYIILATPIAATIIIPPLGLIYFFVORFYVASSRQOLKRLSVSRSPYSHFN 1156
Db 1025 SLFVNTGACIYIILATPIAATIIIPPLGLIYFFVORFYVASSRQOLKRLSVSRSPYSHFN 1084
Qy 1157 ETLGLGVSVIRAFEBEOERFIHOSDKVDENOKAYPSIIVANRWLAIRLECVANGCIVLFPAL 1216
Db 1085 ETLGLGVSVIRAFEBEOERFIHOSDKVDENOKAYPSIIVANRWLAIRLECVANGCIVLFPAL 1144
Qy 1217 PAVISRHSLSAGLVGLSVSYSLQVTTYANMLVRNSESMEETNIIVAVERLKEYSETEKEAPM 1276
Db 1145 PAVISRHSLSAGLVGLSVSYSLQVTTYANMLVRNSESMEETNIIVAVERLKEYSETEKEAPM 1204
Qy 1277 QIOETAPSSWPOVGRVFRNYCLARYEDLPFVLRHIVNTLNGCKKVIVIRGTAGKSSL 1336
Db 1205 QIOETAPSSWPOVGRVFRNYCLARYEDLPFVLRHIVNTLNGCKKVIVIRGTAGKSSL 1264
Qy 1337 TLGAFRINESAGEIIIDIGINIAKIGLHDLRFKIIIPQDPVLFSGSILRMLNDPFSQYSD 1396
Db 1265 TLGAFRINESAGEIIIDIGINIAKIGLHDLRFKIIIPQDPVLFSGSILRMLNDPFSQYSD 1324
Qy 1397 BEVWTSLELAKDFVSLPKDLDHCECAEGGENISVSGQRLVCLARALLRKTILVILDEA 1456
Db 1325 BEVWTSLELAKDFVSLPKDLDHCECAEGGENISVSGQRLVCLARALLRKTILVILDEA 1384

Qy 1457 TAAVDETDLLIOSTIRTOPEDCVTLTIAHRLTIMDYREVIVLDKGEIOEGAPSDLLQ 1516
Db 1385 TAAVDETDLLIOSTIRTOPEDCVTLTIAHRLTIMDYREVIVLDKGEIOEGAPSDLLQ 1444
Qy 1517 ORGLFYSMAVDAGLV 1531
Db 1445 ORGLFYSMAVDAGLV 1459
RESULT 6
ID Q9UQAO_HUMAN PRELIMINARY; PRT; 1456 AA.
Q9UQAO;
AC Q9UQAO;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Multidrug resistance protein (Fragment).
GN Name=MRP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
NCBI_TaxID=9606;
CX [1]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;
RA Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;
RT "Analysis of the intron-exon organization of the human multidrug-
RT resistance protein gene (MRP) and alternative splicing of its mRNA.";
RL Genomics 45:368-378(1997).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF022853; AAB83981.1; -; Genomic DNA.
DR EMBL; AF022824; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022825; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022826; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022827; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022829; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022831; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022833; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022835; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022837; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022847; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022846; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022845; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022844; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022843; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022842; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022841; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022840; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022838; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022852; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022851; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022850; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022849; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022848; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022836; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022834; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022832; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022830; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022828; AAB83981.1; JOINED; Genomic DNA.
DR HSSP; P08716; IMTO.
DR Ensemble; ENSG00000103222; Homo sapiens.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transport.
DR InterPro; IPR003439; ABC_cranep_like.

DR InterPro: IPR005292; MRP assoc.
 DR InterPro: IPR007119; Prot. kinase.
 DR Pfam: PF00664; ABC membrane; 2.
 DR Pfam: PF00005; ABC_tran; 2.
 DR ProDom: PD000006; ABC_transporter; 2.
 DR SMART: SM00382; AAA; 2.
 DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
 DR PROSITE: PS00929; ABC_TM1F; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
 DR PROSITE: PS00893; ABC_TRANSPORTER_2; 2.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR ATP-binding; Nucleotide-binding.
 FT NON_TER 1
 SQ SEQUENCE 1456 AA; 163233 MW; 8DB8AB2BC481F2 CRC64;

Query Match 76.2%; Score 7419.5; DB 2; Length 1456;
 Best Local Similarity 96.1%; Pred. No. 0;
 Matches 1456; Conservative 0; Mismatches 0; Indels 59; Gaps 1;

QY 17 DMNVTMTSNDPDKCFONTVLVWPCPYLMACFPFYLYSRHRCYIQMTPLNKTKTA 76
 DB 1 DMNVTMTSNDPDKCFONTVLVWPCPYLMACFPFYLYSRHRCYIQMTPLNKTKTA 60

QY 77 LGFLMIIVCWADLFYSFWEBSRGIFLAPVPLVSPTLIGITTLATFLIOLERRKGVSSG 136
 DB 61 LGFLMIIVCWADLFYSFWEBSRGIFLAPVPLVSPTLIGITTLATFLIOLERRKGVSSG 120

QY 137 IMTLFMVLVYCALAIIRSKMTMLKEDAOVDLFRDITFYVYFSLIIQLVLCFSRSP 196
 DB 121 IMTLFMVLVYCALAIIRSKMTMLKEDAOVDLFRDITFYVYFSLIIQLVLCFSRSP 180

QY 197 LFSETIHDPNCPBESSASPLSRITFMWITGLIVGYROPLEGSPMLNKEDETSQVVPV 256
 DB 181 LFSETIHDPNCPBESSASPLSRITFMWITGLIVGYROPLEGSPMLNKEDETSQVVPV 240

QY 257 LVNWKKECAKTRKQPKVYVSSKDPQPKSSKVDANEVEALIVSPKEMNPISLFKV 316
 DB 241 LVNWKKECAKTRKQPKVYVSSKDPQPKSSKVDANEVEALIVSPKEMNPISLFKV 300

QY 317 LYKTFGEYFLMSFFFKAIHDLNMFSGPQILKLIKPNYDTKAPDMQGYFYTVLLFVTACL 376
 DB 301 LYKTFGEYFLMSFFFKAIHDLNMFSGPQILKLIKPNYDTKAPDMQGYFYTVLLFVTACL 360

QY 377 QTLVLHGYFHICFVSGRIKTAIVTGAAYRKALVTNSARKSSVYGEIVNLSMDAQRMD 436
 DB 361 QTLVLHGYFHICFVSGRIKTAIVTGAAYRKALVTNSARKSSVYGEIVNLSMDAQRMD 420

QY 437 LATYINMWSAPLOVILALYLLMLNLGSPYLAGVAWVWLVYVNAVWAKTKTYQVAAHK 496
 DB 421 LATYINMWSAPLOVILALYLLMLNLGSPYLAGVAWVWLVYVNAVWAKTKTYQVAAHK 480

QY 497 SKDNRIKLANEILNGIKLVKLIVAMELAFKDKVLAIRQEBELKVKKSAVLSAVGTFTWCT 556
 DB 481 SKDNRIKLANEILNGIKLVKLIVAMELAFKDKVLAIRQEBELKVKKSAVLSAVGTFTWCT 540

QY 557 PELVALCTFAVYVYTDENNILLDAQAVSLAPNLLRFPNLLPMTVSSIVQASVSLKRL 616
 DB 541 PELVALCTFAVYVYTDENNILLDAQAVSLAPNLLRFPNLLPMTVSSIVQASVSLKRL 600

QY 617 RIFLSHELEPDSIERRPVKGGTNSITVRNATFTWARSDDPTLNGITISIEGALVAV 676
 DB 601 RIFLSHELEPDSIERRPVKGGTNSITVRNATFTWARSDDPTLNGITISIEGALVAV 660

QY 677 VGVGVGCKSSLLSALLAEMDKVEGHVAIKGSVAVPOQAMIQNDLSRENILLFCGQLEEPY 736
 DB 661 VGVGVGCKSSLLSALLAEMDKVEGHVAIKGSVAVPOQAMIQNDLSRENILLFCGQLEEPY 688

QY 737 YRSVIOACALLPDLLEILPSGDRTEIGKGVNLSSGQKQRYSLARAVYSNADIVLEDDPLS 796
 DB 689 YRSVIOACALLPDLLEILPSGDRTEIGKGVNLSSGQKQRYSLARAVYSNADIVLEDDPLS 721

QY 797 AYDAHVGHKIFENVIGPKMLKNKTRILLVTHSMYSYLPQVDVITVMSGKISMSGYOELL 856
 DB 797 AYDAHVGHKIFENVIGPKMLKNKTRILLVTHSMYSYLPQVDVITVMSGKISMSGYOELL 856

DB 722 AYDAHVGHKIFENVIGPKMLKNKTRILLVTHSMYSYLPQVDVITVMSGKISMSGYOELL 781
 QY 857 ARDGAFAEPLRTASTQEDDAENNGYTVGSGRPKAKOMENGMVYDTSAGKOLQROLSS 916
 DB 782 ARDGAFAEPLRTASTQEDDAENNGYTVGSGRPKAKOMENGMVYDTSAGKOLQROLSS 841

QY 917 SSSYSGDISHHNSTAELQAKAEKERTWMLMEADKQOTQOVLSYVWMDYKAGLFLSF 976
 DB 842 SSSYSGDISHHNSTAELQAKAEKERTWMLMEADKQOTQOVLSYVWMDYKAGLFLSF 901

QY 977 LSIFLFMCNVSALASNYWLSLWTDDEIVNGTOEHTKVLRSYVGALGISGIAVFGSMA 1036
 DB 902 LSIFLFMCNVSALASNYWLSLWTDDEIVNGTOEHTKVLRSYVGALGISGIAVFGSMA 961

QY 1037 VSTIGLIASRCLHYDLHSLIRSPMSFEPTPSGNLVNRSKELDTYDSMIPEYIKMPMG 1096
 DB 962 VSTIGLIASRCLHYDLHSLIRSPMSFEPTPSGNLVNRSKELDTYDSMIPEYIKMPMG 1021

QY 1097 SLFNVIGACIVILLATPIAIIIPILGLIYEFVQRFVASSROLKRLSVSRSPVSHFN 1156
 DB 1022 SLFNVIGACIVILLATPIAIIIPILGLIYEFVQRFVASSROLKRLSVSRSPVSHFN 1081

QY 1157 ETLIGSVIRAFPEOERFIHQSDLKVDENQKAYVPSIVANRWLAVRLECVGNCIVLFAAL 1216
 DB 1082 ETLIGSVIRAFPEOERFIHQSDLKVDENQKAYVPSIVANRWLAVRLECVGNCIVLFAAL 1141

QY 1217 PAVISRHSLSAGLVGSVSYSLQVTTYLNLVMSSEMETNIYAVERLKEYSETKEAPW 1276
 DB 1142 PAVISRHSLSAGLVGSVSYSLQVTTYLNLVMSSEMETNIYAVERLKEYSETKEAPW 1201

QY 1277 QIOETAPSSWPQVGRFEFPNVCILARYREDJDFPLRHNTVINGEKVGIYGRGAGSSL 1336
 DB 1202 QIOETAPSSWPQVGRFEFPNVCILARYREDJDFPLRHNTVINGEKVGIYGRGAGSSL 1261

QY 1337 TLGLFRINESAGEEIIIDGINIAKIGHDLRFKITTIIPODPVLFSSGLRNULPFSQYSD 1396
 DB 1262 TLGLFRINESAGEEIIIDGINIAKIGHDLRFKITTIIPODPVLFSSGLRNULPFSQYSD 1321

QY 1397 EBYWTSLELAHLKDPVSAIPDKLDHECABEGENLSVGORQVCLARALLRKTILVDEA 1456
 DB 1322 EBYWTSLELAHLKDPVSAIPDKLDHECABEGENLSVGORQVCLARALLRKTILVDEA 1381

QY 1457 TAAVDLETDLLIOSTIRTOEDCTVLTIAHRLNTIDRYVYLDGELTOEGYAPSDLQ 1516
 DB 1382 TAAVDLETDLLIOSTIRTOEDCTVLTIAHRLNTIDRYVYLDGELTOEGYAPSDLQ 1441

QY 1517 QRGLEFSMAKDAGLV 1531
 DB 1442 QRGLEFSMAKDAGLV 1456

RESULT 7
 06UR05 CANFA
 ID 06UR05 CANFA PRELIMINARY; PRT; 1531 AA.
 AC 06UR05;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 GN Multidrug resistance-associated protein 1.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 OC Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22404678; PubMed=12516967;
 RA Ma L., Pratt S.E., Cao J., Dantzig A.H., Moore R.E., Slapak C.A.;
 RT "Identification and characterization of the canine multidrug
 resistance-associated protein."
 RL Mol. Cancer Ther. 1:1335-1342(2002).
 RN [2]

RE NUCLEOTIDE SEQUENCE.
RA Ma L., Prate S.E., Cao J., Danczig A.H., Moore R.E., Slapac C.A.;
RL Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL: AY633728; AAC03148.1; -; mRNA.
DR EMBL: ENSCAFG00000018208; Canis familiaris.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0016887; F:ATPase activity; IEA.
DR GO: GO:0042656; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:0000166; F:nucleotide binding; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR003593; AAA ATPase.
DR InterPro: IPR011527; ABC membrane 1.
DR InterPro: IPR001140; ABC TM transport.
DR InterPro: IPR003439; ABC transp_like.
DR InterPro: IPR005292; MRP assoc.
DR InterPro: IPR000719; Prot_kinase.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD000006; ABC transporter; 2.
DR SMART: SM00382; AAA; 2.
DR TIGRFAMs: TIGR00957; MRP assoc_pro; 1.
DR PROSITE: PS00929; ABC_TM1F; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE: PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE: PS0107; PROTEIN KINASE_ATP; UNKNOWN 1.
DR ATP-binding: Nucleotide-binding; Repeat; Transport.
KW SEQUENCE 1531 AA; 171792 MW; 746361A71C6158BD CRC64;

Query Match 74.7%; Score 7272; DB 2; Length 1531;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 1402; Conservative 70; Mismatches 59; Indels 0; Gaps 0;

QY 1 MALRGFSADSDSLMDMNTMTNSPDTFCQNTVTVWPCRYLMACEFYEYLXSRH 60
DB 1 MALRGFSADSDSLMDMNTMTNSPDTFCQNTVTVWPCRYLMACEFYEYLXSRH 60
QY 61 DRGIQMTPLNKTALGFLIMVCMADLFYSFWEBSRGIFLAPVFLVSPFLIGITLLA 120
DB 61 DRGIQMTPLNKTALGFLIMVCMADLFYSFWEBSRGIFLAPVFLVSPFLIGITLLA 120
QY 121 TFLIQLEBRKGVSSGIMLTFWLVALCALILRSKIMTALKEDAQVDLFEDITFYYVS 180
DB 121 TFLIQLEBRKGVSSGIMLTFWLVALCALILRSKIMTALKEDAQVDLFEDITFYYVS 180
QY 181 LLLIQVLVSCFSDSPFLFSETIHDNPPCESSASTLSITTFWMTGLIVRGYRQPLESD 240
DB 181 LLLIQVLVSCFSDSPFLFSETIHDNPPCESSASTLSITTFWMTGLIVRGYRQPLESD 240
QY 241 LMSLINKEDTSQQVVPVLVKNMKKCAKTRKQPVKVVVSSKDPAPOKESSKYDANEVEAL 300
DB 241 LMSLINKEDTSQQVVPVLVKNMKKCAKTRKQPVKVVVSSKDPAPOKESSKYDANEVEAL 300
QY 301 IVKSPQKEMWNSLFKVLVYKTFEPYFLMSFFKAIHDLMMFSGPOLIKLIFVNDYAPD 360
DB 301 IVKSPQKEMWNSLFKVLVYKTFEPYFLMSFFKAIHDLMMFSGPOLIKLIFVNDYAPD 360
QY 361 WQGFYVYLVFVTAQLQTLVHOYFHTCFVSGMRKTAIVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYVYLVFVTAQLQTLVHOYFHTCFVSGMRKTAIVIGAVYRKALVITNSARKSTV 420
QY 421 GEIYVNLMSVDQRFMDLATYINMIWSAPLOYIALYLLMLNLGSPVLGAVAVMTLVN 480
DB 421 GEIYVNLMSVDQRFMDLATYINMIWSAPLOYIALYLLMLNLGSPVLGAVAVMTLVN 480
QY 481 AVMAKTKTYQVAHMKSKDRIKLMNEILNGIKVLKYAMELAFKQVLAIROBELKYLK 540
DB 481 AVMAKTKTYQVAHMKSKDRIKLMNEILNGIKVLKYAMELAFKQVLAIROBELKYLK 540
QY 541 KSAIYLSAVGTTCVTPFLVALCTFAVYVYIDENNIIDQAQTAFFSLAFNLIRPLNLP 600
DB 541 KSAIYLSAVGTTCVTPFLVALCTFAVYVYIDENNIIDQAQTAFFSLAFNLIRPLNLP 600

QY 601 MVISSIVQASVSLKRLIFLSHELEBDSIERPPVXDGGGTSITVNAATFTWARSDEPT 660
DB 601 MVISSIVQASVSLKRLIFLSHELEBDSIERPPVXDGGGTSITVNAATFTWARSDEPT 660
QY 661 LNGITTSIPGALVAVVGVQCGCKSSLSLLALBMDKVEGHVAIKGSVAVYVQAMQND 720
DB 661 LSGITTSIPGASVAVVGVQCGCKSSLSLLALBMDKVEGHVAIKGSVAVYVQAMQND 720
QY 721 SLRENILFGQLEPPRYRYVIOACALLPDEIILPSGDRTEIGEGKVLSSGQKORVSLAR 780
DB 721 SLRENILFGQLEPPRYRYVIOACALLPDEIILPSGDRTEIGEGKVLSSGQKORVSLAR 780
QY 781 AVYSNADIVLEDDPLSAVDHVGKHIPENYIGPKGMLKNKTRILVTHSMGYLPQVDVIY 840
DB 781 AVYCDSDIYLFDDPLSAVDHVGKHIPENYIGPKGMLKNKTRILVTHSIYLPQVDVIY 840
QY 841 MSGKISSEMSYQELARDAFAFLRTYASTEOQDAENGVTGVSGPKKAKOMNGM 900
DB 841 MTGKISSEMSYQELARDAFAFLRTYASTEOQDAENGVTGVSGPKKAKOMNGM 900
QY 901 LVTDASGKOLOROLSSSSYSGDISRHNSSTAELQKAKEEFTWKLMEADKAQOTGVKL 960
DB 901 LVTDVAGKOLOROLSSSSYSGDVSRHNTSTAEIQKAGPKEDAMKLVADKQOTGVKL 960
QY 961 SVYDYMKAIGLIFSLIFLPMCNHVSALASNYWLSLWTDPIVNGTOEHTKRLSVYG 1020
DB 961 SVYDYMKAIGLIFSLIFLPMCNHVSALASNYWLSLWTDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISQGIANFVGSMAVSIIGIILASRCLHYDLHSILRSMSFFERPPSGNLVNRSEKL 1080
DB 1021 ALGISQGITFVGSMAVSIIGIFASRLHYDLQNVYASPMSPFERPPSGNLVNRSEKL 1080
QY 1081 PTYDSMIPVYIKMFMGSLFNVIAGCIYTLATPIAIIIPPLGIYFFVORFYVASSROL 1140
DB 1081 PTYDSMIPVYIKMFMGSLFNVIAGCIYTLATPIAIIIPPLGIYFFVORFYVASSROL 1140
QY 1141 KRLESVRSVYSHFNFTLLGVSVIRAFEOBREIHQSDLKVDENOKAYPSIVANRMLA 1200
DB 1141 KRLESVRSVYSHFNFTLLGVSVIRAFEOBREIRQSDLKVDENOKAYPSIVANRMLA 1200
QY 1201 VRLCEVNCIVLPALFAVYSRHSLSAGLVGSYSVLOTTYIYNMLVRMSSEMETIVA 1260
DB 1201 VRLCEVNCIVLPALFAVYSRHSLSAGLVGSYSVLOTTYIYNMLVRMSSEMETIVA 1260
QY 1261 VERLKEYSETEKAPWQIOETAPPSWPOYGRVFRYCYLRYPEDLPVLRHINTVINGG 1320
DB 1261 VERLKEYSETEKAPWQIOETAPPSWPOYGRVFRYCYLRYPEDLPVLRHINTVINGG 1320
QY 1321 EKVGIYGRTAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIFQDEVLV 1380
DB 1321 EKVGIYGRTAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIFQDEVLV 1380
QY 1381 SSGLRNMLDPFSQYSDDEEWTSLSLAHLKQFVSALPKLDHECAGEENISVGOQOLVCL 1440
DB 1381 SSGLRNMLDPFSQYSDDEEWTSLSLAHLKQFVSALPKLDHECAGEENISVGOQOLVCL 1440
QY 1441 ARAILRRTKTLIVDEAFAVLDLTDLLOSTIRQFEDCVLTIAHRLNTIMDTRIVL 1500
DB 1441 ARAILRRTKTLIVDEAFAVLDLTDLLOSTIRQFEDCVLTIAHRLNTIMDTRIVL 1500
QY 1501 DKGEIQEYGAPSDILOQRGLFYSMAKDAGLV 1531
DB 1501 DKGEIRCGQPSDILQQRGLFYMAKDAGLV 1531

RESULT 8
059G19 HUMAN
ID 059G19 HUMAN PRELIMINARY; PRT; 1439 AA.
AC 059G19
DT 10-MAY-2005 (Tremblrel. 30, Created)
DT 10-MAY-2005 (Tremblrel. 30, Last sequence update)
DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)

DE ATP-binding cassette, sub-family C, member 1 isoform 1 variant
DE (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain.
RA Tokoh Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title."
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AB209120; BAD92357.1; -; mRNA.
DR Ensemble; ENSG00000103222; Homo sapiens.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC membrane 1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_TM_transp.
DR InterPro; IPR005292; MRP_assoc.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding.
FT NON_TER 1 1
SQ SEQUENCE 1439 AA; 160439 MW; 090D881ADBFB3077 CRC64;
Query Match 74.4%; Score 7242; DB 2; Length 1439;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1426; Conservative 0; Mismatches 1; Indels 10; Gaps 1;
QY 103 APVFLVSPPTLLGTTLLATFLIQLERRKGVSSGIMLTFMLVALVCALAILRSKIMTALK 162
DB 1 APVFLVSPPTLLGTTLLATFLIQLERRKGVSSGIMLTFMLVALVCALAILRSKIMTALK 60
QY 163 EDAGVDLFDITTFVYFSSLLIQVLVSCFSDRSPLFSETIHDNPPCESSASFLSRTTFW 222
DB 61 EDAGVDLFDITTFVYFSSLLIQVLVSCFSDRSPLFSETIHDNPPCESSASFLSRTTFW 120
QY 223 WITGLVIRGYRQPLEGSDLSINKEDTSEGVVPLVKNMKKECAKTRKOPKVVYSSKDP 282
DB 121 WITGLVIRGYRQPLEGSDLSINKEDTSEGVVPLVKNMKKECAKTRKOPKVVYSSKDP 180
QY 283 AQPKESSKVDANEEVEALLVYSPQKEMNPSLFKVLVYTFPGPYFLMSFFFAKIHDLMMFSG 342
DB 181 AQPKESSKVDANEEVEALLVYSPQKEMNPSLFKVLVYTFPGPYFLMSFFFAKIHDLMMFSG 240
QY 343 POLIKLILKFNVDNPKADMGCFYFTVLLFTTACIQTLVHQYFHICVSGMRITKAVIGA 402
DB 241 POLIKLILKFNVDNPKADMGCFYFTVLLFTTACIQTLVHQYFHICVSGMRITKAVIGA 300
QY 403 VYRKALVITNSARKSSTVGEIVNLMSYDAQRFMDLATYINMWSAPQVILALYLMNL 462
DB 301 VYRKALVITNSARKSSTVGEIVNLMSYDAQRFMDLATYINMWSAPQVILALYLMNL 360
QY 463 GPSVLAVAVMVLVFNVAVMAMKTKTYQVAHMSKDNRIKLANEILNGIKVLKYAMEL 522
DB 361 GPSVLAVAVMVLVFNVAVMAMKTKTYQVAHMSKDNRIKLANEILNGIKVLKYAMEL 420

QY 523 AFKDKVLAIRQEBELKVLKKSAYLSAVGFTWVCTPFLVALCTFAVYVYTIIDENNILDAQTA 582
DB 421 AFKDKVLAIRQEBELKVLKKSAYLSAVGFTWVCTPFLVALCTFAVYVYTIIDENNILDAQTA 480
QY 583 FVSLALFNILRFPNLILPMVYSSIVQASVSLKRLRIFLSHELEPDSIERRPVDDGGTN 642
DB 481 FVSLALFNILRFPNLILPMVYSSIVQASVSLKRLRIFLSHELEPDSIERRPVDDGGTN 540
QY 643 SITVRNATFTWASDDPTLNGITFSIPFGALVAVVGVGGCKSSLALSALLAEMDKVGHV 702
DB 541 SITVRNATFTWASDDPTLNGITFSIPFGALVAVVGVGGCKSSLALSALLAEMDKVGHV 600
QY 703 AIKGSVAVVQQAQIMONDSLRENTLFGCQLBEPPYRSVIOACALLPLELILPSGDRREIG 762
DB 601 AIKGSVAVVQQAQIMONDSLRENTLFGCQLBEPPYRSVIOACALLPLELILPSGDRREIG 660
QY 763 EKVNLISGGQKORVSLARAVYNSADLYLPDDPLSAVDAAHKGKHIFFENVIGPKMLKAKTR 822
DB 661 EKVNLISGGQKORVSLARAVYNSADLYLPDDPLSAVDAAHKGKHIFFENVIGPKMLKAKTR 720
QY 823 ILVTHSNSTYLPQVDVITVMSGGKISEMGSYQELIARDGAFRFLRTYASTEQEDAEEN 881
DB 721 ILVTHSNSTYLPQVDVITVMSGGKISEMGSYQELIARDGAFRFLRTYASTEQEDAEEN 780
QY 882 -----GVTVGSGGKAKOMENGLVTDAGKLOLOLSSSSSYSGDISRHNSTA 932
DB 781 STWDEEEAGVTVGSGGKAKOMENGLVTDAGKLOLOLSSSSSYSGDISRHNSTA 840
QY 933 ELQKAEAKKEETWKLMEADRAQTQVQLSYVMDYMKALIGFISPLSFLFMCHNVSALAS 992
DB 841 ELQKAEAKKEETWKLMEADRAQTQVQLSYVMDYMKALIGFISPLSFLFMCHNVSALAS 900
QY 993 NYWLSMTDPIVNGTQEHTRKRLSYVGALGISGIALVFGYSMAVSGIGILASRCHVDL 1052
DB 901 NYWLSMTDPIVNGTQEHTRKRLSYVGALGISGIALVFGYSMAVSGIGILASRCHVDL 960
QY 1053 LHSILBSPMSFFERTPSGNIWNRFSKELDPVDSNIPEVITKMPGSLFNVIAGCIVILLAT 1112
DB 961 LHSILBSPMSFFERTPSGNIWNRFSKELDPVDSNIPEVITKMPGSLFNVIAGCIVILLAT 1020
QY 1113 PIAAIIIPPLGLIYFFQRFYVASSRQIKRLLESYSRSPVYSHFNETLLGVSIVRAFEQE 1172
DB 1021 PIAAIIIPPLGLIYFFQRFYVASSRQIKRLLESYSRSPVYSHFNETLLGVSIVRAFEQE 1080
QY 1173 RFIHQSDLKVDENQKAYPSIVANRWLAVERLCEVGNCTVLFALFAVISRHSLSAGLVGL 1232
DB 1081 RFIHQSDLKVDENQKAYPSIVANRWLAVERLCEVGNCTVLFALFAVISRHSLSAGLVGL 1140
QY 1233 SVSYSIQVTTYLMMVMSSEMETNIVAVRRLKEYSTEEKAPMOIOETAPPSWPQVGR 1292
DB 1141 SVSYSIQVTTYLMMVMSSEMETNIVAVRRLKEYSTEEKAPMOIOETAPPSWPQVGR 1200
QY 1293 VEFRRNCLARYEDLDLFLRHINVTINGEKVYIGVTRGAKSSLTGLFPIINSABEII 1352
DB 1201 VEFRRNCLARYEDLDLFLRHINVTINGEKVYIGVTRGAKSSLTGLFPIINSABEII 1260
QY 1353 IDGINIAKIGLHDLRFKRTIIPQDPVLFGSGLRMNLDPSQYSDEEVTSLBLAHLKDFV 1412
DB 1261 IDGINIAKIGLHDLRFKRTIIPQDPVLFGSGLRMNLDPSQYSDEEVTSLBLAHLKDFV 1320
QY 1413 SAPPDKLDHCCAGGENSELVSQROQLCLARALLRKTKILVLDATAVADLETDDLIOSTI 1472
DB 1321 SAPPDKLDHCCAGGENSELVSQROQLCLARALLRKTKILVLDATAVADLETDDLIOSTI 1380
QY 1473 RTQFEDCTVLTIAHRLNTIMDYTRVIVLDGGEIOETGAPSDLLQOQGLFYSMAKDGLV 1531
DB 1381 RTQFEDCTVLTIAHRLNTIMDYTRVIVLDGGEIOETGAPSDLLQOQGLFYSMAKDGLV 1439
RESULT 9
Q8HX05_BOVIN
ID Q8HX05_BOVIN PRELIMINARY; PRT; 1530 AA.
AC Q8HX05_

DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Multidrug resistance protein 1.
GN Name=MRP1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22062955; PubMed=12067707; DOI=10.1016/S0014-5793(02)02816-8;
RA Taguchi Y., Saeiki K., Komano T.;
RT "Functional analysis of MRP1 cloned from bovine."
RL FEBS Lett. 521:211-213(2002).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AB082124; BAC15550.1; -; mRNA.
DR HSSP; P08716; 1MT0.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity; coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRfams; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1530 AA; 171666 MW; AAEAF92ED7832703 CRC64;
Query Match 74.3%; Score 7230.5; DB 2; Length 1530;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1388; Conservative 83; Mismatches 59; Indels 1; Gaps 1;
QY 1 MALRGFCGADSDPLMDNMVMTNTSNPDTKCFONTVLVWPCFLMACEPPFLYLSRH 60
DB 1 MALNDPCVDSDLFEMWNVMTNTSNPDTKCFONTVLVWPCFLMACEPPFLYLSHH 60
QY 61 DRGYIOMTPLNKTATLGFLLMIYVADLFYSFMRSRGIFLAPYFLVSPITLIGITTLA 120
DB 61 DRGYIOMTHAKAKALGFLLMIYVADLFYSFMRSRGKLLAPFLVSPITLIGITMLA 120
QY 121 TFLIQLERRKGVSSGIMLTFWLVALVCALALRSKIMTALKEDAVDLFRDITFYVFS 180
DB 121 TFLIQLERRRGVSSGIMLTFWLIALCALALRSKIMTALKEDARVDFRDITFYIFS 180
QY 181 LLILQLVLSGDSRSLFSETHDNPCESSASLSLRTTFWITGLVYGRQLDSGD 240
DB 181 LVLQLVLSGDSRSLFSETHDNPCESSASLSLRTTFWITGLVYGRQLDSGD 240
QY 241 LMSLNKEDTSQVNVVLVKNMKKECAKTRKQPVKVYSSKDPAPCKESSKVDANEVAL 300
DB 241 LMSLNKEDTSQVNVVLVKNMKKECAKTRKQPVKVYSSKDPAPCKESSKVDANEVAL 300
QY 301 IVKSPQKEMNPSLFLVLYKTFPGYFLMSFFFKAIHDLNMFSGPQILKLIKFNVDTKAPD 360
DB 301 IVKSPQKEMNPSLFLVLYKTFPGYFLMSFFFKAIHDLNMFSGPQILKLIKFNVDTKAPD 360
QY 361 WQGYTYTLLPTACLQTLVLAHQYHICFVSGMRITKAVIGAVYKALVYTNSSAKSSIV 420
DB 361 WQGYTYTLLPTACLQTLVLAHQYHICFVSGMRITKAVIGAVYKALVYTNSSAKSSIV 420

QY 421 GEIVNLSVDAQRPMDLATYINMISAPLOVITLALYLLMLNLSGVLAGAAMVLMVFN 480
DB 421 GEIVNLSVDAQRPMDLATYINMISAPLOVITLALYLLMLNLSGVLAGAAMVLMVFN 480
QY 481 AVAMKTKTQVAMKSKNRIRIKLMEILLNGIKYLKIYAMELAKDQVLAIRQSELYLK 540
DB 481 AVAMKTKTQVAMKSKNRIRIKLMEILLNGIKYLKIYAMELAKDQVLAIRQSELYLK 540
QY 541 KSAIYAVGTFPTWCTPFLVALCTFAVYVITDENNIIDACTAFPSLAFNLRPEPLNLP 600
DB 541 KSAIYAVGTFPTWCTPFLVALCTFAVYVITDENNIIDACTAFPSLAFNLRPEPLNLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEELDPSIORRPIKDAGATNSITVKNATFTMARNDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEELDPSIORRPIKDAGATNSITVKNATFTMARNDPT 660
QY 661 LNGITPSIPGALVAVNGVCGGSSILSLALAMDVGEVVALKGSVAVYPOQAMOND 720
DB 661 LHGITPSIPGSLVAVNGVCGGSSILSLALAMDVGEVVALKGSVAVYPOQAMONT 720
QY 721 SLRENIILFGCOLBEPYRSVIOACALPDLEILPSGRTIEGKGNLSGGQKORVSLAR 780
DB 721 SLRENIILFGRLQRRYKAVAEACALPDLEILPSGRTIEGKGNLSGGQKORVSLAR 780
QY 781 AVYSNADITYLPDDLAVDAHVGHIFENVIGPKMLKNKTRILVTHSMSTLPQVDVTV 840
DB 781 AVYCDSDVYLLDDPLSAVDADVGHIFENVIGPKMLKNKTRILVTHSMSTLPQVDVTV 840
QY 841 MSGGKISMGSYQELLARDGAFAPFLRTYASTBEQDAENGVGVSGPGKEAKQEMNGM 900
DB 841 MSGGKISMGSYQELLARDGAFAPFLRTYASTBEQDAENGVGVSGPGKEAKQEMNGM 900
QY 901 LVTSAGKOLOROSSSSSYSGDISRRHNSYAELOKAEKKEFTMKLMEADKAQTQVYL 960
DB 901 LVTTTGAKQOROLSSSSSYSDVSOHTTSYAEIRK-PGPFEEFTMKLMEADKAQTQVYL 960
QY 961 SVYDYMKAIGLFTSLIFLFCMCHNVASALSNYMSLMTDDPIVNGTQHTKVLRSYVG 1020
DB 961 SVYDYMKAIGLFTSLIFLFCMCHNVASALSNYMSLMTDDPIVNGTQHTKVLRSYVG 1020
QY 1021 ALGISQGIYVGYMAVSIIGIILASRCLHVDLHLSILRSPSPFEPKPSGNLVNRFSEL 1080
DB 1021 ALGISQGIYVGYMAVSIIGIILASRCLHVDLHLSILRSPSPFEPKPSGNLVNRFSEL 1080
QY 1081 DTYSMTPEVIMKMGSLFNVIYACIVILLATPIAIIITPLGIYFVQGFYVASSRQL 1140
DB 1081 DTYSMTPEVIMKMGSLFNVIYACIVILLATPIAIIITPLGIYFVQGFYVASSRQL 1140
QY 1141 KRLSVSRSPVYSHNETLLGVSVYIAFEQGERFIHOSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLSVSRSPVYSHNETLLGVSVYIAFEQGERFIHOSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLCEVNCIYLPALPAVVISRHSLSAGLVGLSVYSYQVTTYINMLVRMSEMETNIVA 1260
DB 1201 VRLCEVNCIYLPALPAVVISRHSLSAGLVGLSVYSYQVTTYINMLVRMSEMETNIVA 1260
QY 1261 VERLKEVSETEKEAPWQOETAPPSMPQVRVEFRNYCLYREDLDVLAHINVTING 1320
DB 1261 VERLKEVSETEKEAPWQOETAPPSMPQVRVEFRNYCLYREDLDVLAHINVTING 1320
QY 1321 EKVGIYGRGTGAKSLLTGLFRINESAGEIIDDGINIAKIGLHDLARPKITIIIPDPVLF 1380
DB 1321 EKVGIYGRGTGAKSLLTGLFRINESAGEIIDDGINIAKIGLHDLARPKITIIIPDPVLF 1380
QY 1381 SGLSRMNLDPFSQYSDSEVWTSLELAHKDFVSALPDKLDHECAEGGNSLVGQRLVCL 1440
DB 1381 SGLSRMNLDPFSQYSDSEVWTSLELAHKDFVSALPDKLDHECAEGGNSLVGQRLVCL 1440
QY 1441 ARAILRKTIIVLBATVAVDLETDLLQSTIRIQFQEDCYTLTAHRLNTIMDYRYTVL 1500
DB 1441 ARAILRKTIIVLBATVAVDLETDLLQSTIRIQFQEDCYTLTAHRLNTIMDYRYTVL 1500
QY 1501 DKGEIQEGAPSDLLQGRGLFYMAKQAGLV 1531


```

QY 977 LSIFLPCNHNVSALASINWLSLWTDPIVNGTOEHTKRLSVYALGISOGIAVFGSMA 1036
DB 846 LSIFLPCNHNVSALASINWLSLWTDPIVNGTOEHTKRLSVYALGISOGIAVFGSMA 905
QY 1037 VSIQGIILASRGLAHDLSILSRPMSPEERPPSGNLVNRPEKELDTYDMSIPEYIKPMWG 1096
DB 906 VSIQGIILASRGLAHDLSILSRPMSPEERPPSGNLVNRPEKELDTYDMSIPEYIKPMWG 965
QY 1097 SLFVNIAGCIYLLATPIAIIIPPLGIYFVQRFYVASSRQKLRESVSRSPVYSHFN 1156
DB 966 SLFVNIAGCIYLLATPIAIIIPPLGIYFVQRFYVASSRQKLRESVSRSPVYSHFN 1025
QY 1157 ETLIGSVYIRAFEEQERFIHOSDLKVDENQKAYPSIVANRMLAVRECVGNCIVLFAAL 1216
DB 1026 ETLIGSVYIRAFEEQERFIHOSDLKVDENQKAYPSIVANRMLAVRECVGNCIVLFAAL 1085
QY 1217 FAVISRSLASGLVSVSYSLQVTTYLNMVLVRMSSEMETNIIVAEERLKEYSETEKEAPW 1276
DB 1086 FAVISRSLASGLVSVSYSLQVTTYLNMVLVRMSSEMETNIIVAEERLKEYSETEKEAPW 1145
QY 1277 QIOETAPSSWPQVGRVFRNVCRLYREDLDFVLHINVTNGEKGVIIGRTGAGKSSL 1336
DB 1146 QIOETAPSSWPQVGRVFRNVCRLYREDLDFVLHINVTNGEKGVIIGRTGAGKSSL 1205
QY 1337 TLGLFRINESAGEIIIDGINIAKIGHDLAPKTTIIPQDPVLPSSGLRNMIDPFPSQYSD 1396
DB 1206 TLGLFRINESAGEIIIDGINIAKIGHDLAPKTTIIPQDPVLPSSGLRNMIDPFPSQYSD 1265
QY 1397 EEWNTSLFLAHKQFVSLPKLDHECAGEGENISVQROVLCLARALRTKILVDEA 1456
DB 1266 EEWNTSLFLAHKQFVSLPKLDHECAGEGENISVQROVLCLARALRTKILVDEA 1325
QY 1457 TAAVDLETDDLIQSTIRTOFEDCYLTIARLNTIMDTRYVLDKGEIOEGAPSDILQ 1516
DB 1326 TAAVDLETDDLIQSTIRTOFEDCYLTIARLNTIMDTRYVLDKGEIOEGAPSDILQ 1385
QY 1517 QRGLEFYSMAKDAGLV 1531
DB 1386 QRGLEFYSMAKDAGLV 1400

```

```

RESULT 11
MRP1_MOUSE STANDARD; PRT; 1528 AA.
ID MRP1_MOUSE
AC 035379;
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DB Multidrug resistance-associated protein 1 (ATP-binding cassette sub-
DB family C member 1).
GN Name=Abcc1; Synonyms=Abcc1a, Abcc1b, Mdrap, Mrp;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Muscle;
RX MEDLINE=96251691; PubMed=8649356;
RA Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P.C.,
RA Deeley R.G.;
RT "Structure and expression of the messenger RNA encoding the murine
RT multidrug resistance protein, an ATP-binding cassette transporter.",
RL Mol. Pharmacol. 49:962-971(1996).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Miyazaki I., Oeato N., Saito R., Suzuki H., Yamanaoka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,

```

```

RA Schriul L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbett L.B., Cousins S.,
RA Dalla B., Diegani T.A., Fletcher C.F., Foxrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Giesi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Kogiya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Naashima T., Numata K., Okido T., Pavan W.J., Petrea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempile C.A., Teasdale R.D., Tomita K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita K.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hironaka-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Aikawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
CC -1- FUNCTION: May participate directly in the active transport of
CC drugs into subcellular organelles or influence drug distribution
CC indirectly (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.
CC -1- SIMILARITY: Contains 2 ABC transmembrane type-1 domains.
CC -1- SIMILARITY: Contains 2 ABC transporter domains.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AF022908; AAB80938.1; -; mRNA.
DR EMBL; AK029876; BAC26654.1; -; mRNA.
DR HSSP; P08716; IMTO.
DR Ensembl; ENSMUSG00000023088; Mus musculus.
DR MGI; MGI:102676; Abcc1.
DR GO; GO:0005887; C:integral to plasma membrane; IDA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_Membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Glycoprotein; Nucleotide-binding; Repeat; Transmembrane;
KW transport.
FT TOPO_DOM 1 33 Extracellular (By similarity).
FT TRANSMEM 34 54 1 (By similarity).
FT TOPO_DOM 55 74 Cytoplasmic (By similarity).
FT TRANSMEM 75 95 2 (By similarity).
FT TOPO_DOM 96 100 Extracellular (By similarity).
FT TRANSMEM 101 121 3 (By similarity).
FT TOPO_DOM 122 133 Cytoplasmic (By similarity).
FT TRANSMEM 134 154 4 (By similarity).
FT TOPO_DOM 155 172 Extracellular (By similarity).
FT TRANSMEM 173 193 5 (By similarity).
FT TOPO_DOM 194 317 Cytoplasmic (By similarity).
FT TRANSMEM 318 338 6 (By similarity).
FT TOPO_DOM 339 364 Extracellular (By similarity).

```

FT TRANSMEM 365 385 7 (By similarity).
FT TOPO_DOM 386 441 Cytoplasmic (By similarity).
FT TRANSMEM 442 462 8 (By similarity).
FT TOPO_DOM 463 465 Extracellular (By similarity).
FT TRANSMEM 466 486 9 (By similarity).
FT TOPO_DOM 487 548 Cytoplasmic (By similarity).
FT TRANSMEM 549 569 10 (By similarity).
FT TOPO_DOM 570 591 Extracellular (By similarity).
FT TRANSMEM 592 612 11 (By similarity).
FT TOPO_DOM 613 963 Cytoplasmic (By similarity).
FT TRANSMEM 964 984 12 (By similarity).
FT TOPO_DOM 985 1022 Extracellular (By similarity).
FT TRANSMEM 1023 1043 13 (By similarity).
FT TOPO_DOM 1044 1086 Cytoplasmic (By similarity).
FT TRANSMEM 1087 1107 14 (By similarity).
FT TOPO_DOM 1108 1108 Extracellular (By similarity).
FT TRANSMEM 1109 1129 15 (By similarity).
FT TOPO_DOM 1130 1200 Cytoplasmic (By similarity).
FT TRANSMEM 1201 1221 16 (By similarity).
FT TOPO_DOM 1222 1244 17 (By similarity).
FT TRANSMEM 1245 1528 Cytoplasmic (By similarity).
FT TOPO_DOM 1246 1528 ABC transmembrane type-1.1.
FT DOMAIN 326 609 ABC transmembrane type-1.2.
FT DOMAIN 644 868 ABC transporter 1.
FT DOMAIN 971 1253 ABC transporter 2.
FT DOMAIN 1290 1524 ATP 1 (Potential).
FT NE_BIND 678 685 ATP 2 (Potential).
FT NE_BIND 1324 1331 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 19 19 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1003 1003
SQ SEQUENCE 1528 AA; 171185 MW; 68FD13667D61DBB CRC64;

Query Match 71.9%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRGFCSADGSDPLMDMNTYNTSNPDFTKCFONTLVWVPCGYLMACFPYFLYSRH 60
DB 1 MALRSCSADGSDPLMDMNTYNTSNPDFTKCFONTLVWVPCGYLMACFPYFLYSRH 60
QY 61 DRGYIQMTPLNKTALGFLMTVCWADLFYSFWERSRGI FLAPVFLVSPDLLGITLLA 120
DB 61 DRGYIQMTPLNKTALGFLMTVCWADLFYSFWERSRGI FLAPVFLVSPDLLGITLLA 120
QY 121 TELIOERRKGVSSGIMLTFMVALCALAIRSKMTLAKEDAOVDLPRDTTFYYSR 180
DB 121 TELIOERRKGVSSGIMLTFMVALCALAIRSKMTLAKEDAOVDLPRDTTFYYSR 180
QY 121 TELIOERRKGVSSGIMLTFMVALCALAIRSKMTLAKEDAOVDLPRDTTFYYSR 180
DB 121 TELIOERRKGVSSGIMLTFMVALCALAIRSKMTLAKEDAOVDLPRDTTFYYSR 180
QY 181 LLLIOLVLSGFSRSPLESETIHDNPNCPESSASFLSRITFMWITGILVAGYRQPLEGSD 240
DB 181 LLLIOLVLSGFSRSPLESETIHDNPNCPESSASFLSRITFMWITGILVAGYRQPLEGSD 240
QY 241 LMSLNKEDTSEOVVPLVKNMKKECAKTRKQPYKVVYS- SKDPAQKSSKVDANEVEA 299
DB 241 LMSLNKEDTSEOVVPLVKNMKKECAKTRKQPYKVVYS- SKDPAQKSSKVDANEVEA 299
QY 241 LMSLNKEDTSEOVVPLVKNMKKECAKTRKQPYKVVYS- SKDPAQKSSKVDANEVEA 299
DB 241 LMSLNKEDTSEOVVPLVKNMKKECAKTRKQPYKVVYS- SKDPAQKSSKVDANEVEA 299
QY 300 LIYKSPQKEMNBSLFKYLITFGPYFLMSFFKALHDLMFSGPQIILKIKVNDTKAP 359
DB 300 LIYKSPQKEMNBSLFKYLITFGPYFLMSFFKALHDLMFSGPQIILKIKVNDTKAP 359
QY 301 LIYKSPQKEMNBSLFKYLITFGPYFLMSFFKALHDLMFSGPQIILKIKVNDTKAP 359
DB 301 LIYKSPQKEMNBSLFKYLITFGPYFLMSFFKALHDLMFSGPQIILKIKVNDTKAP 359
QY 360 DMGGFYTYVLTFYTAQLOTVLHGYFHICVSGMRITKAVIGAVYKRAVITTSARSSST 419
DB 360 DMGGFYTYVLTFYTAQLOTVLHGYFHICVSGMRITKAVIGAVYKRAVITTSARSSST 419
QY 361 DMGGFYTYVLTFYTAQLOTVLHGYFHICVSGMRITKAVIGAVYKRAVITTSARSSST 420
DB 361 DMGGFYTYVLTFYTAQLOTVLHGYFHICVSGMRITKAVIGAVYKRAVITTSARSSST 420
QY 420 VGEIIVLMSVDAQRFDMLATYINMISAPLOVILALYLLMLNLGPSYLAGAVMLVMPV 479
DB 420 VGEIIVLMSVDAQRFDMLATYINMISAPLOVILALYLLMLNLGPSYLAGAVMLVMPV 479
QY 421 VGEIIVLMSVDAQRFDMLATYINMISAPLOVILALYLLMLNLGPSYLAGAVMLVMPV 480
DB 421 VGEIIVLMSVDAQRFDMLATYINMISAPLOVILALYLLMLNLGPSYLAGAVMLVMPV 480
QY 480 NAYMAKTKTYQVAHMKSKNRIKLMNELLNGIKTLVAMELAFKDKYALIQEELKYL 539
DB 480 NAYMAKTKTYQVAHMKSKNRIKLMNELLNGIKTLVAMELAFKDKYALIQEELKYL 539
QY 481 NAYMAKTKTYQVAHMKSKNRIKLMNELLNGIKTLVAMELAFKDKYALIQEELKYL 540
DB 481 NAYMAKTKTYQVAHMKSKNRIKLMNELLNGIKTLVAMELAFKDKYALIQEELKYL 540
QY 540 KKSAYLSAVGTFTWVCTPFLVALCTFAVVYITDENNILLDAQTAFAVSLALFNILRFLNLT 599
DB 540 KKSAYLSAVGTFTWVCTPFLVALCTFAVVYITDENNILLDAQTAFAVSLALFNILRFLNLT 599

DB 541 KKSAYLSAVGTFTWVCTPFLVALCTFAVVYITDENNILLDAQTAFAVSLALFNILRFLNLT 600
QY 600 PMYISSIVQASVSLKRLIRITLSHELEPDSIERRPVKDGGGNTSYRNAFTWARSDDP 659
DB 601 PMYISSIVQASVSLKRLIRITLSHELEPDSIERRPVKDGGGNTSYRNAFTWARSDDP 659
QY 660 TLNGITFPIEGALVAVVGVGGKSSLSALLAEMDKVGHVTLKSVAVVPOQAIQON 719
DB 660 TLNGITFPIEGALVAVVGVGGKSSLSALLAEMDKVGHVTLKSVAVVPOQAIQON 719
QY 720 DSIURENITLFCQLEEPYRSVIOACALLPDLLEILPSGDRFEIGKGNVLSGGQKQRSIA 779
DB 720 DSIURENITLFCQLEEPYRSVIOACALLPDLLEILPSGDRFEIGKGNVLSGGQKQRSIA 779
QY 780 RAVYSNADIVLPDPLSAVDAAVGHKHI FENVIGPKMKLKKKTRILVTHMSYLPQVDYII 839
DB 780 RAVYSNADIVLPDPLSAVDAAVGHKHI FENVIGPKMKLKKKTRILVTHMSYLPQVDYII 839
QY 840 VMSGKISEMSYQELLDKRGAEFLRTYANAEQDLASDD--SVSGSGKSKPVENG 896
DB 840 VMSGKISEMSYQELLDKRGAEFLRTYANAEQDLASDD--SVSGSGKSKPVENG 896
QY 900 MLVTDGAKOLQOLSSSSSYSGDISRHHNSTAELOKAEKKEETWKLMEADRAQGTQYK 959
DB 900 MLVTDGAKOLQOLSSSSSYSGDISRHHNSTAELOKAEKKEETWKLMEADRAQGTQYK 959
QY 959 LSYVYDMYKAIQGLFISFLSIFLPMCHNSVSLASNYMLSLWTD--PIVNGTQEHTRKLSV 1018
DB 959 LSYVYDMYKAIQGLFISFLSIFLPMCHNSVSLASNYMLSLWTD--PIVNGTQEHTRKLSV 1018
QY 1019 YGALGISQGIYAVGYMAVSIIGILASRCLVDLHLSILRSPMSFERRTSGULVNRFSK 1078
DB 1019 YGALGISQGIYAVGYMAVSIIGILASRCLVDLHLSILRSPMSFERRTSGULVNRFSK 1078
QY 1078 YGALGISQGIYAVGYMAVSIIGILASRCLVDLHLSILRSPMSFERRTSGULVNRFSK 1078
DB 1078 YGALGISQGIYAVGYMAVSIIGILASRCLVDLHLSILRSPMSFERRTSGULVNRFSK 1078
QY 1079 ELDTVDSMTPEVTKMFGSLFENVIGACIVILATPIAIIIPGLGIYFPVQRPYVASSR 1138
DB 1079 ELDTVDSMTPEVTKMFGSLFENVIGACIVILATPIAIIIPGLGIYFPVQRPYVASSR 1138
QY 1136 QLRLESVSRSRSPYSHFNETLLGVSVIRAEEOERFIHOSDLKVDENOKAYPSIVANRW 1198
DB 1136 QLRLESVSRSRSPYSHFNETLLGVSVIRAEEOERFIHOSDLKVDENOKAYPSIVANRW 1198
QY 1199 LAVRLFCVGNCTYLPAALPAVIRSRHSLSAGLVLSYSISQVTTYLNLVYRSGSEMETNI 1258
DB 1199 LAVRLFCVGNCTYLPAALPAVIRSRHSLSAGLVLSYSISQVTTYLNLVYRSGSEMETNI 1258
QY 1259 VAVERLKEYSETEKEAPMIOETAPPSMPOVGRVFRNYCLARYEDLDLVLRHINVTIN 1318
DB 1259 VAVERLKEYSETEKEAPMIOETAPPSMPOVGRVFRNYCLARYEDLDLVLRHINVTIN 1318
QY 1319 GGEKVGIVRTGAGKSLTLGLFRINESABGEIILIDGINIAKIGLHDLRFKTIIPQDPV 1378
DB 1319 GGEKVGIVRTGAGKSLTLGLFRINESABGEIILIDGINIAKIGLHDLRFKTIIPQDPV 1378
QY 1379 LPSGSLRANLDPDSQYSDSEVWTSLELAHLKQFVSLPDKLDHECAEGGENLSVGRQYV 1438
DB 1379 LPSGSLRANLDPDSQYSDSEVWTSLELAHLKQFVSLPDKLDHECAEGGENLSVGRQYV 1438
QY 1439 CLABALLRRTKTLIVDEATRAVDLETFDDLIQSITRIFQFEDCTVLTAAHRLNTIMDTTRYI 1498
DB 1439 CLABALLRRTKTLIVDEATRAVDLETFDDLIQSITRIFQFEDCTVLTAAHRLNTIMDTTRYI 1498
QY 1499 VLDKGEIOEYGAAPSDLLQORGLFYSSMAKQAGLV 1531
DB 1499 VLDKGEIOEYGAAPSDLLQORGLFYSSMAKQAGLV 1531
QY 1496 VLDKGEVRECGAPSELLQORGLFYSSMAKQAGLV 1528
DB 1496 VLDKGEVRECGAPSELLQORGLFYSSMAKQAGLV 1528
RESULT 12
Q810E4 RAT PRELIMINARY; PRT; 1532 AA.
ID Q810E4 RAT
AC Q810E4
DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE ATP-binding cassette protein C1.
GN Name=Abcc1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley, TISUE=SpLeen.
RA Yabuuchi H., Takayanagi S., Ishikawa T.;
RL Submitted (Feb-2002) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF487549; AAC05437.1; -; mRNA.
DR HSSP; P08716; IMTO.
DR Ensemble; ENSRNOG00000032748; Rattus norvegicus.
DR KGD; J112; Abcc1.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0006166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR001395; Aldol/ket_red.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SMO0382; AAA; 2.
DR TIGRfams; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMIF; 2.
DR PROSITE; PS50211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00063; ALDO-KETO_REDUCTASE_3; UNKNOWN_1.
KM ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1532 AA; 171493 MW; 2E639F63F5A3368 CRC64;

Query Match 71.2%; Score 6932.5; DB 2; Length 1532;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 1339; Conservative 99; Mismatches 92; Indels 3; Gaps 3;

QY 1 MALFGSGADSDPLMDMNTWNTSNPDTKCFQNTLVVWPCFFYLACFPFFLYLSRH 60
DB 1 MALSSFCSSDSDPLMDMNTWNTSNPDTKCFQNTLVVWPCFFYLACFPFFLYLSRH 60
QY 61 DRGIQMTPLNKTATAGFLIMIVCMADLFYSFMRSRGIFLAAPFLVSPITLGIITLLA 120
DB 61 DRGIQMTPLNKTATAGFLIMIVCMADLFYSFMRSRGIFLAAPFLVSPITLGIITLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDLDFEDITFYVFS 180
DB 121 TFLIQLERRKGVSSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDLDFEDITFYVFS 180
QY 121 TFLIQLERRKGVSSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDLDFEDITFYVFS 180
DB 121 TFLIQLERRKGVSSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDLDFEDITFYVFS 180
QY 181 LLLIQVLVSCFSDSPFLSETIHDNPPCPSSASLSTITFMWITGLIVRGYROPLESD 240
DB 181 LLLIQVLVSCFSDSPFLSETIHDNPPCPSSASLSTITFMWITGLIVRGYROPLESD 240
QY 181 LLLIQVLVSCFSDSPFLSETIHDNPPCPSSASLSTITFMWITGLIVRGYROPLESD 240
DB 181 LLLIQVLVSCFSDSPFLSETIHDNPPCPSSASLSTITFMWITGLIVRGYROPLESD 240
QY 241 LMSLNKEDTSQVAVLVVKNMKKECAKTRKQPVKVYVS-SKDPAPKESKVDANEVEA 299
DB 241 LMSLNKEDTSQVAVLVVKNMKKECAKTRKQPVKVYVS-SKDPAPKESKVDANEVEA 299
QY 241 LMSLNKEDTSQVAVLVVKNMKKECAKTRKQPVKVYVS-SKDPAPKESKVDANEVEA 299
DB 241 LMSLNKEDTSQVAVLVVKNMKKECAKTRKQPVKVYVS-SKDPAPKESKVDANEVEA 299
QY 300 LIVSPKQKMPSPFLVYKTFGPFYFLMSFFFKAIHDLAMFSGPOLIKLILKFNVDTRAP 359
DB 300 LIVSPKQKMPSPFLVYKTFGPFYFLMSFFFKAIHDLAMFSGPOLIKLILKFNVDTRAP 359
QY 301 LIVSSHDRDPSLFKVLKTFGPFYFLMSFLYKALHDLAMFAGPEILILLINFNVDREAP 360
DB 301 LIVSSHDRDPSLFKVLKTFGPFYFLMSFLYKALHDLAMFAGPEILILLINFNVDREAP 360
QY 360 DWGQYFTVLLFVTAACLOTVLHQYFHLICFVSGMRKIKTAIVGAVYRKALVITNSARKSST 419
DB 360 DWGQYFTVLLFVTAACLOTVLHQYFHLICFVSGMRKIKTAIVGAVYRKALVITNSARKSST 419

DB 361 DWGQYFTVLLFVTAACLOTVLHQYFHLICFVSGMRKIKTAIVGAVYRKALVITNSARKSST 420
QY 420 VGEIVNLSVDAQRFMDLATYINMWSAPLOVILATLMLNLPSTAGAVANVLMVVPV 479
DB 421 VGEIVNLSVDAQRFMDLATYINMWSAPLOVILATLMLNLPSTAGAVANVLMVVPV 480
QY 480 NAYVAMKTKTYQVAHMKSKNRIKLANEILNGIKVLKYAMELAFKQKVLAIROBELKVL 539
DB 481 NAYVAMKTKTYQVAHMKSKNRIKLANEILNGIKVLKYAMELAFKQKVLAIROBELKVL 540
QY 540 KKSAYLSAVGFTFWVCCPFLVALCTFAVYVYTIENNLIDQOTAFVSLATFNILRFPNLI 599
DB 541 KKSAYLSAVGFTFWVCCPFLVALCTFAVYVYTIENNLIDQOTAFVSLATFNILRFPNLI 600
QY 600 PMVSSIVQSVSLKRLIRFLSHHELEPDSIERRPVQGGGTSITVRNATFTWASDDP 659
DB 601 PMVSSIVQSVSLKRLIRFLSHHELEPDSIERRSIDGGGMSITVKNATFTWASDDP 660
QY 660 TLNGITFPAIDGALVAVGVGCGKSLSLALAEKMDKVGSHVTLKGSVAVVPOQAWION 719
DB 661 TLNGITFPAIDGALVAVGVGCGKSLSLALAEKMDKVGSHVTLKGSVAVVPOQAWION 720
QY 720 DSIURENIFPGQLEPPYRSVIOACALPDLEILPSGDRTEIGKGVNLGGQKQORSIA 779
DB 721 DSIURENIFPGQLEPPYRSVIOACALPDLEILPSGDRTEIGKGVNLGGQKQORSIA 780
QY 780 RAYVSNADIVLFDPLSAVDAAHVKHIFENVIGPKGMKNKTRILVTHSMSTLPQVDVVI 839
DB 781 RAYVSNADIVLFDPLSAVDAAHVKHIFENVIGPKGMKNKTRILVTHSMSTLPQVDVVI 840
QY 840 VMSGKTSSEMGVYELLARDGAFEFRTASTEQEODAEBNVGVSGRQKAKQWENG 899
DB 841 VMSGKTSSEMGVYELLARDGAFEFRTASTEQEODAEBNVGVSGRQKAKQWENG 900
QY 900 MLVTVDSAKOLQRLQSSSSYSGDISRHNSTLELQAKAEKERTWLMLEADKQOTGOVK 959
DB 901 MLVTVDSAKOLQRLQSSSSYSGDISRHNSTLELQAKAEKERTWLMLEADKQOTGOVK 959
QY 959 ILLTVDAVGPQRLQSSSSYSGDISRHNSTLELQAKAEKERTWLMLEADKQOTGOVK 959
DB 959 ILLTVDAVGPQRLQSSSSYSGDISRHNSTLELQAKAEKERTWLMLEADKQOTGOVK 959
QY 960 LSVYWDYMKALIGLIFSLIFLPMCNHVSALASNYWLSLWTD-D-PIVNGTOEHTRVLSV 1018
DB 960 LSVYWDYMKALIGLIFSLIFLPMCNHVSALASNYWLSLWTD-D-PIVNGTOEHTRVLSV 1019
QY 1019 YGALGISQIAVFGYSMAVSIIGLILASRCLHVDLHSLNSPMSFFERTPSGNLVNRFSK 1078
DB 1020 YGALGISQIAVFGYSMAVSIIGLILASRCLHVDLHSLNSPMSFFERTPSGNLVNRFSK 1079
QY 1079 ELDTVDSMIPVILKMGSLFENVIGACIVILLAPIAAIIIPGLIYFPQRPYVASSR 1138
DB 1080 ELDTVDSMIPVILKMGSLFENVIGACIVILLAPIAAIIIPGLIYFPQRPYVASSR 1139
QY 1139 QLRKLESVSRSPVYSHNETLLGVSIVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRW 1198
DB 1140 QLRKLESVSRSPVYSHNETLLGVSIVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRW 1199
QY 1199 LAVRLECVGNCTVLPALPAVISRHSLSAGLVGLSVSYSLQVTTYLAWLVMSSEMETNI 1258
DB 1200 LAVRLECVGNCTVLPALPAVISRHSLSAGLVGLSVSYSLQVTTYLAWLVMSSEMETNI 1259
QY 1259 VAVRLEKEYSTEEAQMOTAPSPSSMPQVGVVEERANCLARYEDLDPLRHINVTIN 1318
DB 1260 VAVRLEKEYSTEEAQMOTAPSPSSMPQVGVVEERANCLARYEDLDPLRHINVTIN 1319
QY 1319 GGEKYGIVGRTGAKSSLTGLFPIINSAGEEIIIDGINIAKIGLHMLRFKTIITIPDDV 1378
DB 1320 GGEKYGIVGRTGAKSSLTGLFPIINSAGEEIIIDGINIAKIGLHMLRFKTIITIPDDV 1379
QY 1379 LFGSGLRNNLDPFQSYSDDEEYVTSLELAHLKDPVSALPDKIDHECAEGENLVGQROLV 1438
DB 1380 LFGSGLRNNLDPFQSYSDDEEYVTSLELAHLKDPVSALPDKIDHECAEGENLVGQROLV 1439
QY 1439 CLARALLRKTKILVLDATAVAVDLETDLLQSTIRTOFEDCTVLTIAHRLNTINDYRVI 1498
DB 1440 CLARALLRKTKILVLDATAVAVDLETDLLQSTIRTOFEDCTVLTIAHRLNTINDYRVI 1499

Qy 1499 VLDKGEIOEYGAAPDLLOQRGLFYSSMAKDAGLV 1531
Db 1500 VLDKGEIRCCAPSELLQQRGVFTSMKDAGLV 1532

RESULT 13
Q8CG09 RAT PRELIMINARY; PRT; 1532 AA.

AC Q8CG09_1
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Multidrug resistance-associated protein 1.
GN Name=Abcc1; Synonyms=Mrp1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
NCBI_Taxid=10116;
RN [1]

NP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RA Yang Z., Li C.S.W., Shen D.D., Ho R.J.Y.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL: AY170916; AACN6532.1; -, mRNA.
DR HSSP: P08716; 1MT0.
DR Ensemble1; ENSRNOG0000032748; Rattus norvegicus.
DR RGD; 3112; Abcc1.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR011440; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR001395; Aldo/ket_red.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR Prodom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS00929; ABC_TMIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00863; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00063; ALDO-KETO REDUCTASE_3; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1532 AA; 171558 MW; C18F3554D85732A CRC64;

Query Match 71.2%; Score 6932.5; DB 2; Length 1532;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 1336; Conservative 100; Mismatches 92; Indels 3; Gaps 3;

Qy 1 MALRGFCADGSDPLMDMNVNTWNTSNPDFKCFQNTLVVWPCFYLMACEPFYLYSRH 60
Db 1 MALRSFSSDSDGSDPLMDMNVNTWNTSNPDFKCFQNTLVVWPCFYLMACEPFYLYSRH 60

Qy 61 DRGIVQNTPLNKTITAGFLIMTYCAGDLFYSFWEBSRGIFLAFVPLVSPLLGITTLLA 120
Db 61 DRGIVQNTPLNKTITAGFLIMTYCAGDLFYSFWEBSRGIFLAFVPLVSPLLGITTLLA 120

Qy 121 TFLIQLEERRKGVSSGIMLTFWMLVALYCALAIISKXIMTALKEBAYDLPDITFYVYFS 180
Db 121 TFLIQLEERRKGVSSGIMLTFWMLVALYCALAIISKXIMTALKEBAYDLPDITFYVYFS 180

Qy 121 TFLIQLEERRKGVSSGIMLTFWMLVALYCALAIISKXIMTALKEBAYDLPDITFYVYFS 180
Db 121 TFLIQLEERRKGVSSGIMLTFWMLVALYCALAIISKXIMTALKEBAYDLPDITFYVYFS 180

Qy 181 LLIQLVLSCFSDRSLPFSSETIHDPNCPRESSASFLSRIITFWWITGLIVGYRQPLGSD 240
Db 181 LLIQLVLSCFSDRSLPFSSETIHDPNCPRESSASFLSRIITFWWITGLIVGYRQPLGSD 240

Qy 181 LVFIQLVLSLSCFSDRSLPFSSETIHDPNCPRESSASFLSRIITFWWITGLIVGYRQPLGSD 240
Db 181 LVFIQLVLSLSCFSDRSLPFSSETIHDPNCPRESSASFLSRIITFWWITGLIVGYRQPLGSD 240

Qy 241 LMSLNKEDTSSQVVPVLVKNMKKECAKTRKQPVKVVYS-SKDPAQPKSSKVDANEVEYA 299
Db 241 LMSLNKEDTSSQVVPVLVKNMKKECAKTRKQPVVVAAPDQPKPKSSOLDVANEVEYA 300

Qy 300 LTVKSPKEMWNPSEFLKVLKYPGPFPLMSFEFFKAIHDLMPFSGPQIKLILKFPVNDKAP 359
Db 301 LTVKSHKDNRPSEFLKVLKYPGPFPLMSFLYKALHDLMPAGBEILELINFVNDRAP 360

Qy 360 DWQGYFTYVLLFVTACIQTLVLAHQYFHI CFVSGKRITKTAIVIGAAYRRALVITNSARSKST 419
Db 361 DWQGYFTYVLLFVTACIQTLVLAHQYFHI CFVSGKRITKTAIVIGAAYRRALVITNSARSKST 420

Qy 420 VGEIVNLMSVDAQEFMLATYINMTNSAPLOVILLATVLMNLGSPVLAGAVAWMLMPV 479
Db 421 VGEIVNLMSVDAQEFMLATYINMTNSAPLOVILLATVLMNLGSPVLAGAVAWMLMPV 480

Qy 480 NAWAMAKTKYQVAMHMSKDNRIKLAMBIINGIVLKLAYAMELAFKDKVLAIROBELKVL 539
Db 481 NAWAMAKTKYQVAMHMSKDNRIKLAMBIINGIVLKLAYAMELAFKDKVLAIROBELKVL 540

Qy 540 KKSAYLSAVGTFTWCTPFLVALCTFAVYVITDENNTLDAQAFVSLAFNILRFPPLNL 599
Db 541 KKSAYLSAVGTFTWCTPFLVALCTFAVYVITDENNTLDAKKAFLVSLAFNILRFPPLNL 600

Qy 600 PMVTSITVOASVSLKRLRILPSHELEPDSIERRPVDDGGTNSITVRNATFTMARSDPP 659
Db 601 PMVTSITVOASVSLKRLRILPSHELEPDSIERRSIXDGGMNSITVKNATFTMARSDPP 660

Qy 660 TLNGITSIPEGALVAVVGVGCGKSSLSLALAEQMKVEGHVAIKGSVAVPOAMQON 719
Db 661 TLNGITSIPEGALVAVVGVGCGKSSLSLALAEQMKVEGHVILKGSVAVPOAMQON 720

Qy 720 DSIENENILFCQLEPEPYRSVIOACALLPDLLEILPSGDRTEIGKGVNLSGGQKQVSLA 779
Db 721 DSIENENILFCQLEPEPYRSVIOACALLPDLLEILPSGDRTEIGKGVNLSGGQKQVSLA 780

Qy 780 RAYVSNADITLFPDDPLSAVDAHVGKAI FENVITGPKGMLKNTKRLVLVTHSMSTLPQVDYII 839
Db 781 RAYVSNADITLFPDDPLSAVDAHVGKAI FENVITGPKGMLKNTKRLVLVTHSMSTLPQVDYII 840

Qy 840 VMSGGKISEMGSVOELIARDAQAPAEFLRTYASTBOEDAEENGVTGVSQPKKAKOMENG 899
Db 841 VMSGGKISEMGSVOELIARDAQAPAEFLRTYASTBOEDAEENGVTGVSQPKKAKOMENG 900

Qy 900 MLVYDSAGKQLOQOLSSSSSSSGDISRHNSSTAELQAKAKKEETWKLMEADKAQTOGVK 959
Db 901 MLVYDSAGKQLOQOLSSSSSSSGDISRHNSSTAELQAKAKKEETWKLMEADKAQTOGVK 959

Qy 960 LSYVWDYMKAIIGLFIPLSLIFLMCNHVSALASNYWLSLMTDD-PIYNGTQEHKTVRLSV 1018
Db 960 LSYVWDYMKAIIGLFIPLSLIFLMCNHVSALASNYWLSLMTDD-PIYNGTQEHKTVRLSV 1019

Qy 1019 YGALGISQGIAGVFGYSAVSIIGLILASRCHLVLLHSILSPMSFEERTSGNLYNFSK 1078
Db 1020 YGALGISQGIAGVFGYSAVSIIGLILASRCHLVLLHSILSPMSFEERTSGNLYNFSK 1079

Qy 1079 ELDTVDSMIPEVAKMFNGSLFNVIACIVILLATPIAIIIPILGLIFYFFQRFYVASSR 1138
Db 1080 ELDTVDSMIPEVAKMFNGSLFNVIACIVILLATPIAIIIPILGLIFYFFQRFYVASSR 1139

Qy 1139 QLRLESVSRSRSPYVSHRNEFTLLGVSVIRABEEDRFTHQSDLYVDENOKAYYSIVANRW 1198
Db 1140 QLRLESVSRSRSPYVSHRNEFTLLGVSVIRABEEDRFTHQSDLYVDENOKAYYSIVANRW 1199

Qy 1199 LAVRLCEVGNCIYLPALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVLRMSSEMTNI 1258
Db 1200 LAVRLCEVGNCIYLPALFAVISRHSLSAGLVGLSVSYSLQVTTYLNLVLRMSSEMTNI 1259

Qy 1259 VAVERLKEYSETEKEAPWQIOETAPSPSPOVGRVVEFRNYCLARYRBDLVFLRHINVTIN 1318
Db 1260 VAVERLKEYSETEKEASWQIOETAPSPSWHSGRVERFRDYCLARYRBDLVFLRHINVTIS 1319

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 15:05:25 ; Search time 39.0811 Seconds
(without alignments)
4076.547 Million cell updates/sec

Title: US-10-665-283-6

Perfect score: 9901

Sequence: 1 MALRGCASGSDPLMDMN.....ASVAVAKRKPSPDLSL 1927

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: Issued Patents AA: *
2: /cgn2_6/ptodata/1/iaa/5 COMB.pdp: *
3: /cgn2_6/ptodata/1/iaa/6 COMB.pdp: *
4: /cgn2_6/ptodata/1/iaa/H_COMB.pdp: *
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pdp: *
6: /cgn2_6/ptodata/1/iaa/backfile1.pdp: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7860	79.4	1531	1 US-08-463-092B-4	Sequence 4, Appli
2	7860	79.4	1531	1 US-08-462-109A-4	Sequence 4, Appli
3	7860	79.4	1531	1 US-08-460-907B-4	Sequence 4, Appli
4	7860	79.4	1531	2 US-08-463-179A-4	Sequence 4, Appli
5	7860	79.4	1531	2 US-08-461-384B-4	Sequence 4, Appli
6	7860	79.4	1531	2 US-09-647-140B-19	Sequence 19, Appli
7	7849	79.3	1531	1 US-08-141-893-2	Sequence 2, Appli
8	7849	79.3	1531	1 US-08-463-092B-2	Sequence 2, Appli
9	7849	79.3	1531	1 US-08-462-109A-2	Sequence 2, Appli
10	7849	79.3	1531	1 US-08-460-907B-2	Sequence 2, Appli
11	7849	79.3	1531	2 US-08-463-179A-2	Sequence 2, Appli
12	7849	79.3	1531	2 US-08-461-384B-2	Sequence 2, Appli
13	7849	79.3	1531	2 US-08-407-207A-2	Sequence 2, Appli
14	7002.5	70.7	1528	1 US-08-463-092B-6	Sequence 6, Appli
15	7002.5	70.7	1528	1 US-08-462-109A-6	Sequence 6, Appli
16	7002.5	70.7	1528	1 US-08-460-907B-6	Sequence 6, Appli
17	7002.5	70.7	1528	2 US-08-463-179A-6	Sequence 6, Appli
18	7002.5	70.7	1528	2 US-08-461-384B-6	Sequence 6, Appli
19	4487.5	45.3	1527	2 US-09-647-140B-6	Sequence 6, Appli
20	4487.5	45.3	1530	2 US-09-647-140B-33	Sequence 33, Appli
21	3391.5	34.3	1503	2 US-09-647-140B-8	Sequence 8, Appli
22	3385.5	34.2	1503	2 US-09-792-616-3	Sequence 3, Appli
23	3218.5	32.5	1498	2 US-09-792-616-9	Sequence 3, Appli
24	2374.5	24.0	1621	2 US-08-972-927-3	Sequence 3, Appli
25	2349.5	23.7	1622	2 US-08-972-927-6	Sequence 6, Appli
26	2266.5	22.9	1325	2 US-09-647-140B-2	Sequence 2, Appli
27	2212.5	22.3	1464	2 US-10-012-896-1008	Sequence 1008, Ap

28	2195.5	22.2	1261	2 US-09-439-313-538	Sequence 538, App
29	2195.5	22.2	1261	2 US-09-636-215-538	Sequence 538, App
30	2195.5	22.2	1261	2 US-08-685-166A-538	Sequence 538, App
31	2195.5	22.2	1261	2 US-09-679-426-538	Sequence 538, App
32	2195.5	22.2	1261	2 US-09-759-143-538	Sequence 538, App
33	2195.5	22.2	1261	2 US-09-651-236-538	Sequence 538, App
34	2195.5	22.2	1261	2 US-09-657-279-538	Sequence 538, App
35	2195.5	22.2	1261	2 US-10-012-896-538	Sequence 538, App
36	2195.5	22.2	1261	2 US-10-012-896-1009	Sequence 1009, App
37	2118	21.4	1581	2 US-08-726-320-3	Sequence 3, Appli
38	2118	21.4	1581	2 US-09-208-716-3	Sequence 3, Appli
39	2115	21.4	1228	2 US-09-439-313-537	Sequence 537, App
40	2115	21.4	1228	2 US-09-636-215-537	Sequence 537, App
41	2115	21.4	1228	2 US-09-685-166A-537	Sequence 537, App
42	2115	21.4	1228	2 US-09-679-426-537	Sequence 537, App
43	2115	21.4	1228	2 US-09-759-143-537	Sequence 537, App
44	2115	21.4	1228	2 US-09-651-236-537	Sequence 537, App
45	2115	21.4	1228	2 US-09-657-279-537	Sequence 537, App

ALIGNMENTS

RESULT 1
US-08-463-092B-4
Sequence 4, Application US/08463092B
Patent No. 5766880
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
MULTIDRUG RESISTANCE PROTEINS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,092B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: Q1546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-092B-4

```

```

Query Match      79.4%; Score 7860; DB 1; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MALRGFCSADGSDPLMDMNTWNTSNPFTKCFQNTVLVWVPCRYLWACFFPYLYLSRH 60
DB 1 MALRGFCSADGSDPLMDMNTWNTSNPFTKCFQNTVLVWVPCRYLWACFFPYLYLSRH 60
QY 61 DRGYIQTPLNKTALGFLIMIVCMADLFYSFMRSGRIFLAPVFLVSPILLGITLLA 120
DB 61 DRGYIQTPLNKTALGFLIMIVCMADLFYSFMRSGRIFLAPVFLVSPILLGITLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYS 180
DB 121 TFLIQLERRKGVSSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYS 180
QY 181 LLLIQLVLSGSDSPFLFSETIHDNPPCRESSASFLSKITFWMTTGLIVRGYRQPLBESD 240
DB 181 LLLIQLVLSGSDSPFLFSETIHDNPPCRESSASFLSKITFWMTTGLIVRGYRQPLBESD 240
QY 241 LMSLNKEDTSBOQVPLVVKMKKECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVEAL 300
DB 241 LMSLNKEDTSBOQVPLVVKMKKECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVEAL 300
QY 301 IVKSPKEMNPSLFKVLYKTFGPFYFLMSFFKAIHDLMMFSGPOLIKLLIFVNDTKAPD 360
DB 301 IVKSPKEMNPSLFKVLYKTFGPFYFLMSFFKAIHDLMMFSGPOLIKLLIFVNDTKAPD 360
QY 361 WQGFYVYVLFVTAQLQVLVHOYFHICFVSGMRKKTAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYVYVLFVTAQLQVLVHOYFHICFVSGMRKKTAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIYNLMSVDKQRFMDLATYINMIWSAPLOYILALYLLMLNGPSVLGAVVWMLMVPVN 480
DB 421 GEIYNLMSVDKQRFMDLATYINMIWSAPLOYILALYLLMLNGPSVLGAVVWMLMVPVN 480
QY 481 AVMAKTKTYOAVAMKSKDNRIKLMNEILLNGIKVLKLYAMELAFKQVLAIROBELKYLK 540
DB 481 AVMAKTKTYOAVAMKSKDNRIKLMNEILLNGIKVLKLYAMELAFKQVLAIROBELKYLK 540
QY 541 KSAVLSAVGTFTWCTPPLVALCTPAVYVITDENNILLDAQTAFFSLAFNILLRPLNLTLP 600
DB 541 KSAVLSAVGTFTWCTPPLVALCTPAVYVITDENNILLDAQTAFFSLAFNILLRPLNLTLP 600
QY 601 MVISIVQASVSLKRLRIFLSHEBELPDSIRRPVKDGGNSITVRNATFTWARSDDPT 660
DB 601 MVISIVQASVSLKRLRIFLSHEBELPDSIRRPVKDGGNSITVRNATFTWARSDDPT 660
QY 661 LMGITFSPREGALVAVVGQVCGKSSLLSALLAEMDKYEGHVALIKGSVAVYVPOQAMIOND 720
DB 661 LMGITFSPREGALVAVVGQVCGKSSLLSALLAEMDKYEGHVALIKGSVAVYVPOQAMIOND 720
QY 721 SLRENIIIFGCOLLEBPYRVSVOACALLPDLBILPSGDRTEIGEKCUNISGGOKORVSLAR 780
DB 721 SLRENIIIFGCOLLEBPYRVSVOACALLPDLBILPSGDRTEIGEKCUNISGGOKORVSLAR 780
QY 781 AVYNSNADLYLPDDPLSAVDAAVGHKIFENNVIGPKMKLNKRIILVTHSMVSLPOVDVITV 840
DB 781 AVYNSNADLYLPDDPLSAVDAAVGHKIFENNVIGPKMKLNKRIILVTHSMVSLPOVDVITV 840
QY 841 MSGKISMSGSYOBELADGAFABELRTYASTEQDADAEENGVTGVSFGKEAKOMENGM 900
DB 841 MSGKISMSGSYOBELADGAFABELRTYASTEQDADAEENGVTGVSFGKEAKOMENGM 900
QY 901 LVTDSAGKOLQOROLSSSSSYGDISRHNSITAELOKAEKKEETWKLMEADKAQTGYVL 960
DB 901 LVTDSAGKOLQOROLSSSSSYGDISRHNSITAELOKAEKKEETWKLMEADKAQTGYVL 960

```

```

DB 901 LVTDSAGKOLQOROLSSSSSYGDISRHNSITAELOKAEKKEETWKLMEADKAQTGYVL 960
QY 961 SVYWDYMKAIQGLFISFISIFLFMCNHYSAASNTWLSLMTDDPIVNGTOEHTKYRLSVYG 1020
DB 961 SVYWDYMKAIQGLFISFISIFLFMCNHYSAASNTWLSLMTDDPIVNGTOEHTKYRLSVYG 1020
QY 1021 ALGISQGIAVFGYSMAVSIQGIILASRCLHYDLHLSILRSPMSFFERTPSGNLVNRFSEL 1080
DB 1021 ALGISQGIAVFGYSMAVSIQGIILASRCLHYDLHLSILRSPMSFFERTPSGNLVNRFSEL 1080
QY 1081 DTVDSMTPEVYKMGMSLFPNVIGACIVILATPILAIILIPPLGIYFVQRFYVASSROL 1140
DB 1081 DTVDSMTPEVYKMGMSLFPNVIGACIVILATPILAIILIPPLGIYFVQRFYVASSROL 1140
QY 1141 KRLESVRSPPYSHFNSTLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLESVRSPPYSHFNSTLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLECVNCCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETNIVA 1260
DB 1201 VRLECVNCCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMGOIETAPSSMPOYGRVFERRYCLRYREDLPVLRHINVTINGG 1320
DB 1261 VERLKEYSETEKEAPMGOIETAPSSMPOYGRVFERRYCLRYREDLPVLRHINVTINGG 1320
QY 1321 EKVGIVGRTGAGKSLTLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTLIIPQDPVLF 1380
DB 1321 EKVGIVGRTGAGKSLTLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTLIIPQDPVLF 1380
QY 1381 SGLRPNULDPPSOYSDEEVTWSTELAHLKDFVSALPKLDHECAEGEENISVGOROLVCL 1440
DB 1381 SGLRPNULDPPSOYSDEEVTWSTELAHLKDFVSALPKLDHECAEGEENISVGOROLVCL 1440
QY 1441 ARALLRKTILIVDEAFAVADLETDLLQSTIRQFEDCTVLTAAHRLNTIMDTRYIVL 1500
DB 1441 ARALLRKTILIVDEAFAVADLETDLLQSTIRQFEDCTVLTAAHRLNTIMDTRYIVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531

RESULT 2
US-08-462-109A-4
; Sequence 4, Application US/08462109A
; Patent No. 5882875
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: LAHIVE & COCKFIELD
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,109A
; FILING DATE:
; CLASSIFICATION: 435
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340

```

FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: POI-002CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-4

Query Match 79.4%; Score 7860; DB 1; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFSADGSDPLMDMNTMTNTPDFTKCPONTVLVWPCFYLMACPFYFLYSRH 60
DB 1 MALRGFSADGSDPLMDMNTMTNTPDFTKCPONTVLVWPCFYLMACPFYFLYSRH 60
QY 61 DRGVIQMTPLNKTKTALGFLMIVCMADLFYSFWEBSRGIFLAVFLVSPTLGITTLLA 120
DB 61 DRGVIQMTPLNKTKTALGFLMIVCMADLFYSFWEBSRGIFLAVFLVSPTLGITTLLA 120
QY 121 TPLQLERRKGVSSGIMLTFMLVALCALAIIILSKITMTAKEDAOVDFRDTITFYVFS 180
DB 121 TPLQLERRKGVSSGIMLTFMLVALCALAIIILSKITMTAKEDAOVDFRDTITFYVFS 180
QY 121 TPLQLERRKGVSSGIMLTFMLVALCALAIIILSKITMTAKEDAOVDFRDTITFYVFS 180
DB 121 TPLQLERRKGVSSGIMLTFMLVALCALAIIILSKITMTAKEDAOVDFRDTITFYVFS 180
QY 181 LLLIQLVLSGFSRDSPLFSETHDNPNCRESSASFLSRITPMTITGLIVGYRQLGSD 240
DB 181 LLLIQLVLSGFSRDSPLFSETHDNPNCRESSASFLSRITPMTITGLIVGYRQLGSD 240
QY 241 LMSLNKEDTSBOVPLVVKWKKCAKTRKQPVKVVYSKDPAPKESKVDANEVEAL 300
DB 241 LMSLNKEDTSBOVPLVVKWKKCAKTRKQPVKVVYSKDPAPKESKVDANEVEAL 300
QY 301 IVVSPQKEMNPSPFKVLYKTGPFLMSFPFKAIHDLMTSGPQIILKIKFVNDTAPD 360
DB 301 IVVSPQKEMNPSPFKVLYKTGPFLMSFPFKAIHDLMTSGPQIILKIKFVNDTAPD 360
QY 361 MOGFYFVLLFVTACLOTVLVHOVFNHCPVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420
DB 361 MOGFYFVLLFVTACLOTVLVHOVFNHCPVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420
QY 421 GEIYNLMSVDAOREMDLATYINMIWSAPLQVITLALYLLMINTLGPVSLAGVAVMLAMPVN 480
DB 421 GEIYNLMSVDAOREMDLATYINMIWSAPLQVITLALYLLMINTLGPVSLAGVAVMLAMPVN 480
QY 481 AVNMAKTYTYQVAMHMSKDNRIKIMNEILNGIKVLYKLYAMELAFKDXVLAIROBELVKLK 540
DB 481 AVNMAKTYTYQVAMHMSKDNRIKIMNEILNGIKVLYKLYAMELAFKDXVLAIROBELVKLK 540
QY 541 KSAVLSAVGTFTWCTPELVALCTPAVYVITDENNIIIDAQTAPEVSLANILRPPLNLP 600
DB 541 KSAVLSAVGTFTWCTPELVALCTPAVYVITDENNIIIDAQTAPEVSLANILRPPLNLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPVKOGGNTSITVNAATFTWARSDDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPVKOGGNTSITVNAATFTWARSDDPT 660
QY 661 LNTGTFSPREGALVAVVGVGGCGSSLSLALAMDVEYGHVAKGSVAVYPOAMQND 720
DB 661 LNTGTFSPREGALVAVVGVGGCGSSLSLALAMDVEYGHVAKGSVAVYPOAMQND 720
QY 721 SLRENILFGQLBEPYRSVIAQCALLPDLIELLPSGDRTEIGEGVNLSSGQKQVSLAR 780

DB 721 SLRENILFGQLBEPYRSVIAQCALLPDLIELLPSGDRTEIGEGVNLSSGQKQVSLAR 780
QY 781 AVYSNADIYLPDDPLSADVDAHVGKHIFENYIGPKMKLKNKTRILVTHSMSTLPQVDYIIV 840
DB 781 AVYSNADIYLPDDPLSADVDAHVGKHIFENYIGPKMKLKNKTRILVTHSMSTLPQVDYIIV 840
QY 841 MSGKLSKEMSGYELLARDGAPAEFLRTVASTOEOEAENGVGVSGPKKAKOMNGM 900
DB 841 MSGKLSKEMSGYELLARDGAPAEFLRTVASTOEOEAENGVGVSGPKKAKOMNGM 900
QY 901 LVYDSAGKQORQLSSSSSYSGDISRRHNSHTAELOKAEKETKLMKADKQTOGVKL 960
DB 901 LVYDSAGKQORQLSSSSSYSGDISRRHNSHTAELOKAEKETKLMKADKQTOGVKL 960
QY 961 SVYWDYKKAIGLIFSLIFSLFMCNHSALASNYWLSLMTDDPIVNGTOHTKVLRSVYG 1020
DB 961 SVYWDYKKAIGLIFSLIFSLFMCNHSALASNYWLSLMTDDPIVNGTOHTKVLRSVYG 1020
QY 1021 ALGISOGIAVGYSMAYSIGIILASRCLHYDLHSLIRSPMSFEERTPSGNLVNRSKEL 1080
DB 1021 ALGISOGIAVGYSMAYSIGIILASRCLHYDLHSLIRSPMSFEERTPSGNLVNRSKEL 1080
QY 1081 DTVDSMIPVYIKMFMSLFNVIGACIYIILATPILAIIPPLGLIYFFVORFYVASSROL 1140
DB 1081 DTVDSMIPVYIKMFMSLFNVIGACIYIILATPILAIIPPLGLIYFFVORFYVASSROL 1140
QY 1141 KRLSVRSRSPVSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENQKAYPSIVANWLA 1200
DB 1141 KRLSVRSRSPVSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENQKAYPSIVANWLA 1200
QY 1201 VRLCVCNCIVLPAALPAVISRHSLSAGLVGSVSVSLQVTTYLNTLVRSSEMETNIVA 1260
DB 1201 VRLCVCNCIVLPAALPAVISRHSLSAGLVGSVSVSLQVTTYLNTLVRSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMOIQETAPSSWPQVGRVEFRNYCLRYBEDLDFVLRHINVTNGG 1320
DB 1261 VERLKEYSETEKAPMOIQETAPSSWPQVGRVEFRNYCLRYBEDLDFVLRHINVTNGG 1320
QY 1321 EKVGIWGTAGKSSLTGLFRINESABEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
DB 1321 EKVGIWGTAGKSSLTGLFRINESABEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
QY 1381 SSGSLRNLDPFSQYSDBEWTSLFLAKDPFVSLPKLDHECAEGENISVGOQLVCL 1440
DB 1381 SSGSLRNLDPFSQYSDBEWTSLFLAKDPFVSLPKLDHECAEGENISVGOQLVCL 1440
QY 1441 ARALLRKTKLIVDEATAVADLETDDLIQSTIRTOFEDCTVLIHRLNNTIMDYTRYIVL 1500
DB 1441 ARALLRKTKLIVDEATAVADLETDDLIQSTIRTOFEDCTVLIHRLNNTIMDYTRYIVL 1500
QY 1501 DKGEIOEYGA PSDLLQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGA PSDLLQORGLFYSMAKDAGLV 1531

RESULT 3
US-08-460-907B-4
Sequence 4, Application US/08460907B
Patent No. 5891724
GENERAL INFORMATION:
APPLICANT: Deeley, Roger G.
APPLICANT: Cole, Susan P.C.
TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG
RESISTANCE ON A CELL
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: PARTDO RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,907B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01551
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-907B-4

Query Match      79.4%; Score 7860; DB 1; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MALRGFCAGDSDPLMDNNTWNTSNPDFTKCFQNTVLVWVPCFYLMACFPFFYLYLSRH 60
DB      1 MALRGFCAGDSDPLMDNNTWNTSNPDFTKCFQNTVLVWVPCFYLMACFPFFYLYLSRH 60
QY      61 DRGYIQMTPLNKTKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPTLGITTLLA 120
DB      61 DRGYIQMTPLNKTKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPTLGITTLLA 120
QY      121 TFLIQLERRKVVQSSGIMLTFWLVALVCALALIRKIKTALKEDQVNDLFRDITTYVVS 180
DB      121 TFLIQLERRKVVQSSGIMLTFWLVALVCALALIRKIKTALKEDQVNDLFRDITTYVVS 180
QY      121 TFLIQLERRKVVQSSGIMLTFWLVALVCALALIRKIKTALKEDQVNDLFRDITTYVVS 180
DB      121 TFLIQLERRKVVQSSGIMLTFWLVALVCALALIRKIKTALKEDQVNDLFRDITTYVVS 180
QY      181 LLLIQLVSCSDSPLESETIHDNPPCBSSASFLSITTFWIMIGLIVRGYROPLEGGSD 240
DB      181 LLLIQLVSCSDSPLESETIHDNPPCBSSASFLSITTFWIMIGLIVRGYROPLEGGSD 240
QY      241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKQPVKVVYSSKDPAPQKSSKVDANEVEAL 300
DB      241 LMSLNKEDTSQVVPVLVKNMKKECAKTRKQPVKVVYSSKDPAPQKSSKVDANEVEAL 300
QY      301 IVKSPQKWNNSLPFVLYKTGPFYFLMSFFPKAIDLMFSGPOLIKLIKRVNDTKAPD 360
DB      301 IVKSPQKWNNSLPFVLYKTGPFYFLMSFFPKAIDLMFSGPOLIKLIKRVNDTKAPD 360
QY      361 WQGFYTYLTFYTAQLQTLVLAHQYFHCIVSGMRKTKAVTGAIVRKALVITNSAKKSTV 420
DB      361 WQGFYTYLTFYTAQLQTLVLAHQYFHCIVSGMRKTKAVTGAIVRKALVITNSAKKSTV 420
QY      421 GEIVNLSVDAQRFMDLATYINIMWSAPLQVTLALYLLMLNIGPSVLGAVAVMLVPEVN 480
```

```

DB      421 GEIVNLSVDAQRFMDLATYINIMWSAPLQVTLALYLLMLNIGPSVLGAVAVMLVPEVN 480
QY      481 AVAMAKTKTYQVAHMKSKDNRIKLMNEILINGIKVLKLYAMELAFKDKVLAIRQELKYLK 540
DB      481 AVAMAKTKTYQVAHMKSKDNRIKLMNEILINGIKVLKLYAMELAFKDKVLAIRQELKYLK 540
QY      541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNILDQAOTAVSVALFNILRPLNLP 600
DB      541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNILDQAOTAVSVALFNILRPLNLP 600
QY      601 MVISSTIVQASVSLKRLAIFLSHELEPDSIERRVKQCGGNSITVNAATFTWASDPPT 660
DB      601 MVISSTIVQASVSLKRLAIFLSHELEPDSIERRVKQCGGNSITVNAATFTWASDPPT 660
QY      661 LINGITFSPREGALVAVVGVQCGGKSLSLALAMDKYEGHVAIKGSVAAYVPOQAMTQND 720
DB      661 LINGITFSPREGALVAVVGVQCGGKSLSLALAMDKYEGHVAIKGSVAAYVPOQAMTQND 720
QY      721 SLRENILFGCQLEBPYRSVYIQACALLPDLLEILPSGDRTEIGKGVNLSSGQKORVSLAR 780
DB      721 SLRENILFGCQLEBPYRSVYIQACALLPDLLEILPSGDRTEIGKGVNLSSGQKORVSLAR 780
QY      781 AVYSNADIYLFDDPLSAVDHVGKHIPEENVYGPXGMLKNKTRILLVTHSMSTLPQVDVTV 840
DB      781 AVYSNADIYLFDDPLSAVDHVGKHIPEENVYGPXGMLKNKTRILLVTHSMSTLPQVDVTV 840
QY      841 MSGGKISEMGSYQELRLARDGAFAEFLRTYASTEOQDAEENGVTGVSQPGKEAKQEMNGM 900
DB      841 MSGGKISEMGSYQELRLARDGAFAEFLRTYASTEOQDAEENGVTGVSQPGKEAKQEMNGM 900
QY      901 LVYDSAGKQLOROLSSSSSYSGDISRRHNSYAELOKAEXKEETWKLMEADKAOTGOYKL 960
DB      901 LVYDSAGKQLOROLSSSSSYSGDISRRHNSYAELOKAEXKEETWKLMEADKAOTGOYKL 960
QY      961 SVYWDYMKALGLFISFLSIFLMCNHVSALASNWMLSLMTDDPLVNGQETHTKRLSYVG 1020
DB      961 SVYWDYMKALGLFISFLSIFLMCNHVSALASNWMLSLMTDDPLVNGQETHTKRLSYVG 1020
QY      1021 ALGISOGIAVFGYMAVSIIGIILASRCIHDVLLSILRSPMSFFERTPSGULVNRFSKEL 1080
DB      1021 ALGISOGIAVFGYMAVSIIGIILASRCIHDVLLSILRSPMSFFERTPSGULVNRFSKEL 1080
QY      1081 DTVDSMIPEVIKMGMSLFNVIGACIVLLATPAAIIIPPLGIYFFVQGFYVASSRQL 1140
DB      1081 DTVDSMIPEVIKMGMSLFNVIGACIVLLATPAAIIIPPLGIYFFVQGFYVASSRQL 1140
QY      1141 KRLSVSRSPYVSHFNETLLGVSVITRAPEQERTIHQSDDLKVDENOKAYYSIVANRWLA 1200
DB      1141 KRLSVSRSPYVSHFNETLLGVSVITRAPEQERTIHQSDDLKVDENOKAYYSIVANRWLA 1200
QY      1201 VRLCEVGNCTVLPALPAFVVISRHSLSAGLVGLSVYSIQVTTYVNLVVRMSSEMETNIVA 1260
DB      1201 VRLCEVGNCTVLPALPAFVVISRHSLSAGLVGLSVYSIQVTTYVNLVVRMSSEMETNIVA 1260
QY      1261 VERLKEYSETEKEAPWQIOETAPSPSSPOVGRVFFRNVCYLRBDLPVLRHINVTINGG 1320
DB      1261 VERLKEYSETEKEAPWQIOETAPSPSSPOVGRVFFRNVCYLRBDLPVLRHINVTINGG 1320
QY      1321 EKVGIVGRTGAKGSLTGLFRINESAGEIITIDGINIAKIGLHDLRKRTIYIIPQDPVLF 1380
DB      1321 EKVGIVGRTGAKGSLTGLFRINESAGEIITIDGINIAKIGLHDLRKRTIYIIPQDPVLF 1380
QY      1381 SGLSRMNLDPFSQVSDSEVWTSLELAHKDVSVALPDGLDHECARGENLSVGRQVQLCL 1440
DB      1381 SGLSRMNLDPFSQVSDSEVWTSLELAHKDVSVALPDGLDHECARGENLSVGRQVQLCL 1440
QY      1441 ARALLRKTKIIVLDEATAAVDELFTDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVVL 1500
DB      1441 ARALLRKTKIIVLDEATAAVDELFTDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVVL 1500
QY      1501 DKGEIOEYGAQSDLLIQRGFLFYSAKDAGLV 1531
```

Db 1501 DKGEIOEYCAPSDLLQORGLFYSMAKADGLV 1531

RESULT 4
US-08-463-179A-4
Sequence 4, Application US/08463179A
Patent No. 6001563
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,179A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Deconcti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: POI-002CP8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-179A-4

Query Match 79.4%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFGSADGSDPLMDMNTWNTSNDFYKCFONTVLVWVPCFYLMACFPFYLYLSRH 60
Db 1 MALRGFGSADGSDPLMDMNTWNTSNDFYKCFONTVLVWVPCFYLMACFPFYLYLSRH 60
QY 61 DRGVIQMTPLNKTALGFLIMIVCWADLFYSPWERSGIFLAVFLVSPFLTGITTLA 120
Db 61 DRGVIQMTPLNKTALGFLIMIVCWADLFYSPWERSGIFLAVFLVSPFLTGITTLA 120
QY 121 TFLIQLERRKGVSSGIMLTFMLVALVCAILRSKIMTALKEADQVDFEDITFYVFS 180
Db 121 TFLIQLERRKGVSSGIMLTFMLVALVCAILRSKIMTALKEADQVDFEDITFYVFS 180
QY 181 LLILQLVLSGSDSPFLSETIHDPNCPRESSASFLSRTITWTGLIVRGYRQPLRGSD 240
Db 181 LLILQLVLSGSDSPFLSETIHDPNCPRESSASFLSRTITWTGLIVRGYRQPLRGSD 240
QY 241 LMSLNKEDTSBQVAVLVKMKWKECAKTRKQPVKVVYSSKDPAPKSSSKVDANBEVYL 300

Db 241 LMSLNKEDTSBQVAVLVKMKWKECAKTRKQPVKVVYSSKDPAPKSSSKVDANBEVYL 300
QY 301 IVKSPQKEMNPSLTKVLYKTFGPFYLMSPFKAIHDLMMSPGPOILKLLIFVNDTKAPD 360
Db 301 IVKSPQKEMNPSLTKVLYKTFGPFYLMSPFKAIHDLMMSPGPOILKLLIFVNDTKAPD 360
QY 361 MGGYFYTVLLPVTACLOTVLVHOYFHI CFPSGMGIKTAIVGAVYRKALVITNSARKSTV 420
Db 361 MGGYFYTVLLPVTACLOTVLVHOYFHI CFPSGMGIKTAIVGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDAQRPMDLATYINMWSAPLOVIALYLLMLNGSVSLAGVAVMLAMVYN 480
Db 421 GEIVNLMSVDAQRPMDLATYINMWSAPLOVIALYLLMLNGSVSLAGVAVMLAMVYN 480
QY 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLYKLYAMELAFKDVLAIRQELKYLK 540
Db 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLYKLYAMELAFKDVLAIRQELKYLK 540
QY 541 KSAVLAVGFTWCTPFLVALCTFAVYVITDENNILLDAQTA FVSLAFNLRPPLNLP 600
Db 541 KSAVLAVGFTWCTPFLVALCTFAVYVITDENNILLDAQTA FVSLAFNLRPPLNLP 600
QY 601 MVISSIVQASVSLKRLIFLSHEELPDSIERREPVKGGGTSITVNNATFTMARSDPT 660
Db 601 MVISSIVQASVSLKRLIFLSHEELPDSIERREPVKGGGTSITVNNATFTMARSDPT 660
QY 661 LMGITFSIPBGALVAVVGVGCGKSSLSALLAEMDKVEGHVAIKGSVAVVPOQAMIOND 720
Db 661 LMGITFSIPBGALVAVVGVGCGKSSLSALLAEMDKVEGHVAIKGSVAVVPOQAMIOND 720
QY 721 SLRENILFGCOLPEPYRVSIVQACALLPDLFLPSGRTEIGEGVNLSGGQKQVSLAR 780
Db 721 SLRENILFGCOLPEPYRVSIVQACALLPDLFLPSGRTEIGEGVNLSGGQKQVSLAR 780
QY 781 AVVSNADIYLPDDPLSAVDHVGKHIPEENVIGPQMKKNTRIILVTHSMSTLPQVDYIV 840
Db 781 AVVSNADIYLPDDPLSAVDHVGKHIPEENVIGPQMKKNTRIILVTHSMSTLPQVDYIV 840
QY 841 MSGGKISEMGSYQELLARDGAFAEFLRTVASTEBODABENGVTGSGPGKEAKOMENGM 900
Db 841 MSGGKISEMGSYQELLARDGAFAEFLRTVASTEBODABENGVTGSGPGKEAKOMENGM 900
QY 901 LVYDSAKQIQRQULSSSSSTISGDISRHNSHTAEIQKEAKKEETWKLMADKQOTGYKL 960
Db 901 LVYDSAKQIQRQULSSSSSTISGDISRHNSHTAEIQKEAKKEETWKLMADKQOTGYKL 960
QY 961 SVYWDYMKAIGLFISFLSIFLFMCNHYSAIASNYWLSLMTDDPIVNGTOEHTKYRLSYG 1020
Db 961 SVYWDYMKAIGLFISFLSIFLFMCNHYSAIASNYWLSLMTDDPIVNGTOEHTKYRLSYG 1020
QY 1021 ALGISOGIAVFGYMAVSIGAILASRCLAHVLDLSILRSFMSFPERTPSGNLVNRFSKEL 1080
Db 1021 ALGISOGIAVFGYMAVSIGAILASRCLAHVLDLSILRSFMSFPERTPSGNLVNRFSKEL 1080
QY 1081 DTYDSMIPBYIKMFMSLENVIGACIYVILATPIAIIIPPLGLITYFVORFYVASSROL 1140
Db 1081 DTYDSMIPBYIKMFMSLENVIGACIYVILATPIAIIIPPLGLITYFVORFYVASSROL 1140
QY 1141 KRLSVSRSPVYSFNFETLLGVSIVTRAFFEOERFIHOSDLKVDENQAAVYPSIYANMFLA 1200
Db 1141 KRLSVSRSPVYSFNFETLLGVSIVTRAFFEOERFIHOSDLKVDENQAAVYPSIYANMFLA 1200
QY 1201 VRLCEVNCIIVLFAALFAVLSRHSLSAGVLGSYSVLAQVTTYLNMVLVRMSSEMETIVA 1260
Db 1201 VRLCEVNCIIVLFAALFAVLSRHSLSAGVLGSYSVLAQVTTYLNMVLVRMSSEMETIVA 1260
QY 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVFRNYCLRREDDLVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVFRNYCLRREDDLVLRHINVTINGG 1320
QY 1321 EKVGIVGRTAGKSSLLGLGFRINESAGEGIIIDGIIAKIGLHDLRFKTIIPQDVLVLF 1380

Db 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEIITIDGINIAKIGLHDLRFKTTITIPDDVLF 1380
Qy 1381 SGLSRMLNLPFSQYSDSEEVMTSLIHLKDFVSALPDLDBHCAEGENLSVGQRQVCL 1440
Db 1381 SGLSRMLNLPFSQYSDSEEVMTSLIHLKDFVSALPDLDBHCAEGENLSVGQRQVCL 1440
Qy 1441 ARALLRKTIIVLDEATAVDLETDLLIQTIRTOFEDCTVLTIAHRLNTMDYTRIVL 1500
Db 1441 ARALLRKTIIVLDEATAVDLETDLLIQTIRTOFEDCTVLTIAHRLNTMDYTRIVL 1500
Qy 1501 DKGEIOEYGAPSDILQQRGLFYMAKXAGLV 1531
Db 1501 DKGEIOEYGAPSDILQQRGLFYMAKXAGLV 1531
RESULT 5
US-08-461-384B-4
Sequence 4, Application US/08461384B
Patent No. 6023473
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESS: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,384B
FILING DATE: 05-JUN-95
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01547
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-384B-4
Query Match 79.4%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MALRFGCADSGSDPMDNMVNTSNPDFTKCFONTVLVWVPCFYLMACFPFYFLYSRH 60
Db 1 MALRFGCADSGSDPMDNMVNTSNPDFTKCFONTVLVWVPCFYLMACFPFYFLYSRH 60
Qy 61 DRGVIQMTPLNKTKTALGFLMIYCMADLFYSFWERSRGIFLAPVFLVSPILLGITTLIA 120
Db 61 DRGVIQMTPLNKTKTALGFLMIYCMADLFYSFWERSRGIFLAPVFLVSPILLGITTLIA 120

Db 61 DRGVIQMTPLNKTKTALGFLMIYCMADLFYSFWERSRGIFLAPVFLVSPILLGITTLIA 120
Qy 121 TFLIQLERRKGVSSGIMLTFMWALVALCALILRSKIMTALKEDAQVDLFFEDITFYVYS 180
Db 121 TFLIQLERRKGVSSGIMLTFMWALVALCALILRSKIMTALKEDAQVDLFFEDITFYVYS 180
Qy 181 LLLIQLVLSCFSDRSPLEFSETIHDNPNCPSSSASFSLRITWMTTGLIVRGYRQPLESSD 240
Db 181 LLLIQLVLSCFSDRSPLEFSETIHDNPNCPSSSASFSLRITWMTTGLIVRGYRQPLESSD 240
Qy 241 LMSLNKEDTSQVVPVLVYKMKKECAKTRKQPVVYVSSKQPAOKESSKYDANEVVAL 300
Db 241 LMSLNKEDTSQVVPVLVYKMKKECAKTRKQPVVYVSSKQPAOKESSKYDANEVVAL 300
Qy 301 IVKSPKEMNPSLFKVLKTFGPYFLMSFFPKALHDLMMFSGPOLKLLIFVNDYKAPD 360
Db 301 IVKSPKEMNPSLFKVLKTFGPYFLMSFFPKALHDLMMFSGPOLKLLIFVNDYKAPD 360
Qy 361 WQGYFYTVLLFVTKCLQTLVLHGYFHICFVSGMRKTAIVGAVYRKALVITNSARKSTV 420
Db 361 WQGYFYTVLLFVTKCLQTLVLHGYFHICFVSGMRKTAIVGAVYRKALVITNSARKSTV 420
Qy 421 GEIVNLMSVDAQRFMDLATTYINMIMSAPLOVITLALYLLMLAGPSVLGAVAMVLMVYN 480
Db 421 GEIVNLMSVDAQRFMDLATTYINMIMSAPLOVITLALYLLMLAGPSVLGAVAMVLMVYN 480
Qy 481 AVAMAKTKTYOAHMKSKDNRIKLMNEILNGIKVLYKUYAMELAFKDYALROBELKYLK 540
Db 481 AVAMAKTKTYOAHMKSKDNRIKLMNEILNGIKVLYKUYAMELAFKDYALROBELKYLK 540
Qy 541 KSAVLSAVGTTWCTPPLVALCTPNAVYTTIDENNILDAQTAFLSALFNILRPPLNLP 600
Db 541 KSAVLSAVGTTWCTPPLVALCTPNAVYTTIDENNILDAQTAFLSALFNILRPPLNLP 600
Qy 601 MVISIVQASVSLRIRIFLSHEELPDSIRRPVKOGGNSITVRATFTMARSDPT 660
Db 601 MVISIVQASVSLRIRIFLSHEELPDSIRRPVKOGGNSITVRATFTMARSDPT 660
Qy 661 LMGITFSPREGALVAVVQVCGKSSLSALLAEMDKVEGVAIKGSVAVYVPOQAMIOND 720
Db 661 LMGITFSPREGALVAVVQVCGKSSLSALLAEMDKVEGVAIKGSVAVYVPOQAMIOND 720
Qy 721 SLRENILFGCOLLEBPYRSVIOACALLPDLILPSGDRTEIGEKGVNLSGGOKORVSLAR 780
Db 721 SLRENILFGCOLLEBPYRSVIOACALLPDLILPSGDRTEIGEKGVNLSGGOKORVSLAR 780
Qy 781 AVYSNADIVLFPDDPLSANDAHVGHIFENVIGPKMLKNKTRILVTHSMSTLPOVDVITV 840
Db 781 AVYSNADIVLFPDDPLSANDAHVGHIFENVIGPKMLKNKTRILVTHSMSTLPOVDVITV 840
Qy 841 MSGKISEMSGYOELLARDGAFAEFLRTYASTBOQDAEENGVTGVSQPGKEAKOMENGM 900
Db 841 MSGKISEMSGYOELLARDGAFAEFLRTYASTBOQDAEENGVTGVSQPGKEAKOMENGM 900
Qy 901 LVYDSAGKOLOROLSSSSSYSGDISRRHNSYAELOKAEAKKEETWKLMEADAKOTGYVL 960
Db 901 LVYDSAGKOLOROLSSSSSYSGDISRRHNSYAELOKAEAKKEETWKLMEADAKOTGYVL 960
Qy 961 SVYDYMKAIGLFIISFLSIFLMCHNVASALSNWLSLMTDPIYNGOEHTKVRLSYG 1020
Db 961 SVYDYMKAIGLFIISFLSIFLMCHNVASALSNWLSLMTDPIYNGOEHTKVRLSYG 1020
Qy 1021 ALGISOGIAVFGYSMAVSIIGLILASRCHADVLDSILSPMSFFERTSGNLYNFSKEL 1080
Db 1021 ALGISOGIAVFGYSMAVSIIGLILASRCHADVLDSILSPMSFFERTSGNLYNFSKEL 1080
Qy 1081 DTVDSMIBEVLKMFWSLENVIGACIVILLATPINAIIIPGLGIYFFVQRPYVASSQL 1140
Db 1081 DTVDSMIBEVLKMFWSLENVIGACIVILLATPINAIIIPGLGIYFFVQRPYVASSQL 1140
Qy 1141 KRLSVSSSPYVSHNETLLGSVYTRAFEBQRFPHQSDLVKDEQKAYRYSIVANRWLA 1200
Db 1141 KRLSVSSSPYVSHNETLLGSVYTRAFEBQRFPHQSDLVKDEQKAYRYSIVANRWLA 1200

QY 1201 VRLECVGNCIVLFAALFAVSRHSLSAGLVG.SVYSLSQVTTYLNMLVRMSSEMETNIVA 1260
Db 1201 VRLECVGNCIVLFAALFAVSRHSLSAGLVG.SVYSLSQVTTYLNMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFRNYCLARYREDLDFVLRHINVTNGG 1320
Db 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFRNYCLARYREDLDFVLRHINVTNGG 1320
QY 1321 EKVIVTGTGAGKSLTGLFRINESAGEIINGINAKIGLHDLRPKITITIPQDPVLF 1380
Db 1321 EKVIVTGTGAGKSLTGLFRINESAGEIINGINAKIGLHDLRPKITITIPQDPVLF 1380
QY 1381 SGRSRMNDPSPQSDSEEWTSLELAHKDFVSALPDKLDEHCAEGGENLSVGOROLVCL 1440
Db 1381 SGRSRMNDPSPQSDSEEWTSLELAHKDFVSALPDKLDEHCAEGGENLSVGOROLVCL 1440
QY 1441 ABALLRKTILVDEATAVADLETDDLIQSTIRTOFEDCTVLTAAHRLNTIMDYTRVIL 1500
Db 1441 ABALLRKTILVDEATAVADLETDDLIQSTIRTOFEDCTVLTAAHRLNTIMDYTRVIL 1500
QY 1501 DKGEIOEYGAPSDILQORGLFYSMADAGLV 1531
Db 1501 DKGEIOEYGAPSDILQORGLFYSMADAGLV 1531

RESULT 6

US-09-647-140B-19
; Sequence 19, Application US/09647140B
; Patent No. 6803184
; GENERAL INFORMATION:
; APPLICANT: Fox Chase Cancer Center
; APPLICANT: Kruth, Gary D.
; APPLICANT: Lee, Kun
; APPLICANT: Belinsky, Martin G.
; APPLICANT: Bain, Lisa J.
; TITLE OF INVENTION: MRP-Related ABC Transporter Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use Thereof
; FILE REFERENCE: PCCC 98-02
; CURRENT APPLICATION NUMBER: US/09/647,140B
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US99/06644
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/079,759
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/095,153
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-647-140B-19

Query Match 79.4%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCADGSDPLDMANTYNTSNDFTKCFQNTLVVWPCFYLACFPFYLYSRH 60
Db 1 MALRGFCADGSDPLDMANTYNTSNDFTKCFQNTLVVWPCFYLACFPFYLYSRH 60
QY 61 DRGYIOMTPIKXTKTAAGLLMIVCMADLFYSPWERSRGIFLAVFLVSPTLGITTLA 120
Db 61 DRGYIOMTPIKXTKTAAGLLMIVCMADLFYSPWERSRGIFLAVFLVSPTLGITTLA 120
QY 121 TPILIOERRKGVSSGIMLFPMVALYCALIISKMTALKRAQVDLRODTFFYYFS 180
Db 121 TPILIOERRKGVSSGIMLFPMVALYCALIISKMTALKRAQVDLRODTFFYYFS 180
QY 181 LLLIQVLVSCFSDRSPFLFSETIHDNPNCPESSASFLSRITFWWITGLIVRGYRPLEGSD 240
Db 181 LLLIQVLVSCFSDRSPFLFSETIHDNPNCPESSASFLSRITFWWITGLIVRGYRPLEGSD 240

Db 181 LLLIQVLVSCFSDRSPFLFSETIHDNPNCPESSASFLSRITFWWITGLIVRGYRPLEGSD 240
QY 241 LMSLNKEDTSBOVVPVYVNRWKKECACTRRQPKYVNVYSSSDPAOPKSSXYDANEVEAL 300
Db 241 LMSLNKEDTSBOVVPVYVNRWKKECACTRRQPKYVNVYSSSDPAOPKSSXYDANEVEAL 300
QY 301 IVKSPQKEMNPSPLEKVLVYKTFGPFYFLMSFFFKAIHDMFMSPQIILKLLIKFVNDTAPD 360
Db 301 IVKSPQKEMNPSPLEKVLVYKTFGPFYFLMSFFFKAIHDMFMSPQIILKLLIKFVNDTAPD 360
QY 361 MOGYFTVLLFVTAACQTLVLAHQYFHCIFVSGMRKIKTAIVAGAVRKALVITNSARKSSTV 420
Db 361 MOGYFTVLLFVTAACQTLVLAHQYFHCIFVSGMRKIKTAIVAGAVRKALVITNSARKSSTV 420
QY 421 GEIVNLMSVDAORPMDLATYINMISAPLOYIILALYILNLGPSVLAGVAVWYLPVN 480
Db 421 GEIVNLMSVDAORPMDLATYINMISAPLOYIILALYILNLGPSVLAGVAVWYLPVN 480
QY 481 AVAMAKTKTYQVAMHMSKDNRIKLMBEILNGIKVLYLAMELAFKDKVLAIROBELKVLK 540
Db 481 AVAMAKTKTYQVAMHMSKDNRIKLMBEILNGIKVLYLAMELAFKDKVLAIROBELKVLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVYTDENNILDAQTAFFVSLAFNILRPLNLP 600
Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVYTDENNILDAQTAFFVSLAFNILRPLNLP 600
QY 601 MVTSSIVQASVLSKRLRIFLSHELEBDSIERRPVKDGGGNTSITVENATFTMARSDPT 660
Db 601 MVTSSIVQASVLSKRLRIFLSHELEBDSIERRPVKDGGGNTSITVENATFTMARSDPT 660
QY 661 LMGITFSIPGALVAVVGVGCGKSSLSLALBMDVBEHVAIKSGVAVYPOQAWOND 720
Db 661 LMGITFSIPGALVAVVGVGCGKSSLSLALBMDVBEHVAIKSGVAVYPOQAWOND 720
QY 721 SLRENILFGQLEBEPYRYSVIOACALPDLEILPSCGRTEIGEGVNLSSGQKORVSLAR 780
Db 721 SLRENILFGQLEBEPYRYSVIOACALPDLEILPSCGRTEIGEGVNLSSGQKORVSLAR 780
QY 781 AVYSNADIVYFDDPLSANDAHVGHIFENVIGPKMKKNKTRILVTHSMGYLPQVDYIIV 840
Db 781 AVYSNADIVYFDDPLSANDAHVGHIFENVIGPKMKKNKTRILVTHSMGYLPQVDYIIV 840
QY 841 MSGGKISEMWSYOBELARDGAFELRTVASTOEOBAEENGVTGSGPKKAKOMNGM 900
Db 841 MSGGKISEMWSYOBELARDGAFELRTVASTOEOBAEENGVTGSGPKKAKOMNGM 900
QY 901 LVYDSAGKQORQLSSSSSYSGDISRHHNSTAELOKAEAKKETWKLMEADKAOQGVKL 960
Db 901 LVYDSAGKQORQLSSSSSYSGDISRHHNSTAELOKAEAKKETWKLMEADKAOQGVKL 960
QY 961 SVYWDYMKALGLFISPLSIFLPMCNHVSALASNYWLSLWTDPIVNGTOGHTYVRLSVYG 1020
Db 961 SVYWDYMKALGLFISPLSIFLPMCNHVSALASNYWLSLWTDPIVNGTOGHTYVRLSVYG 1020
QY 1021 ALGISOGIANFVGSMASVIGGIIASRCLEHVDLHSLIRSMSPFRRPSGNLVNRPEKL 1080
Db 1021 ALGISOGIANFVGSMASVIGGIIASRCLEHVDLHSLIRSMSPFRRPSGNLVNRPEKL 1080
QY 1081 DTVDSMIEPVYIKNFMGSLFNVIGACIYILLATPIAIIIPPLGIYEFVORFYVASSROL 1140
Db 1081 DTVDSMIEPVYIKNFMGSLFNVIGACIYILLATPIAIIIPPLGIYEFVORFYVASSROL 1140
QY 1141 KRLESVRSRSPVYSHFNETLLGVSVIRAFEOERFIIHQSDLKVDENOKAYVPSIVANRWLA 1200
Db 1141 KRLESVRSRSPVYSHFNETLLGVSVIRAFEOERFIIHQSDLKVDENOKAYVPSIVANRWLA 1200
QY 1201 VRLECVGNCIVLFAALFAVSRHSLSAGLVG.SVYSLSQVTTYLNMLVRMSSEMETNIVA 1260
Db 1201 VRLECVGNCIVLFAALFAVSRHSLSAGLVG.SVYSLSQVTTYLNMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFRNYCLARYREDLDFVLRHINVTNGG 1320
Db 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFRNYCLARYREDLDFVLRHINVTNGG 1320

QY 1321 EKVGIVGRTGAGKSLTGLFRINESAGEIITIDGINIAKIGLHDLREKTIITIPQDPVLF 1380
DB 1321 EKVGIVGRTGAGKSLTGLFRINESAGEIITIDGINIAKIGLHDLREKTIITIPQDPVLF 1380
QY 1381 SGSIKRMULDPPSQSDEEWTSLSLAHLKDFVSALPDLDHECAEGGNSLVGGQOLVCL 1440
DB 1381 SGSIKRMULDPPSQSDEEWTSLSLAHLKDFVSALPDLDHECAEGGNSLVGGQOLVCL 1440
QY 1441 ARALRKTKILVLEBATAVDLETDLLQSTIRTOFEDCTVLTJAHRLNTIMDYTRVVL 1500
DB 1441 ARALRKTKILVLEBATAVDLETDLLQSTIRTOFEDCTVLTJAHRLNTIMDYTRVVL 1500
QY 1501 DKGEIOEGAPSDLLQORGLFYMAKQAGLV 1531
DB 1501 DKGEIOEGAPSDLLQORGLFYMAKQAGLV 1531

RESULT 7
US-08-141-893-2
; Sequence 2, Application US/08141893
; Patent No. 5489519
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: MOLITIDRUG RESISTANCE PROTEIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/141,893
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923; 08/029,340
; FILING DATE: 27-OCT-1992; 8-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A. Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: P01-002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5149
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-141-893-2

Query Match 79.3%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCSDGSDPLMDMNVNTWNTSNBDFTKCFONTVLVWPCFYLMACFPFFYLYLSRH 60
DB 1 MALRGFCSDGSDPLMDMNVNTWNTSNBDFTKCFONTVLVWPCFYLMACFPFFYLYLSRH 60
QY 61 DRGIQWTPMLNKTATAGFLIMIVCMADLFYSFWRSGIGIFLAPFVLSPTLLGTTTLA 120
DB 61 DRGIQWTPMLNKTATAGFLIMIVCMADLFYSFWRSGIGIFLAPFVLSPTLLGTTTLA 120

QY 121 TFLIQLERRKGVSSGIMLTFWLVALVCALAILRSKTMALKEDAQVDLFDITFYVFS 180
DB 121 TFLIQLERRKGVSSGIMLTFWLVALVCALAILRSKTMALKEDAQVDLFDITFYVFS 180
QY 181 LLLIQVLSCGSDPSPLFSETHDNPCEPSSASFLSRITFWMTTGLIVGGRPLGSD 240
DB 181 LLLIQVLSCGSDPSPLFSETHDNPCEPSSASFLSRITFWMTTGLIVGGRPLGSD 240
QY 241 LMSLNKEDTSEQVVPVVKWKKCECAKTRKQPVVYVSSKDPAPKESSKVADAEVEAL 300
DB 241 LMSLNKEDTSEQVVPVVKWKKCECAKTRKQPVVYVSSKDPAPKESSKVADAEVEAL 300
QY 301 IVKSPQKEMNPSELFKVLKTYTGPYFLMSPEFKAIHDLMMFSGPOLKLLIKFVNDYAPD 360
DB 301 IVKSPQKEMNPSELFKVLKTYTGPYFLMSPEFKAIHDLMMFSGPOLKLLIKFVNDYAPD 360
QY 361 WQGYFYTVLLFVTAQCTOTLVHOFYHICFVSGMKIKTAIVGAVYRKALVITNSARKSTV 420
DB 361 WQGYFYTVLLFVTAQCTOTLVHOFYHICFVSGMKIKTAIVGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINMIMSAPLOVILALYLLMLNLAGSVLAGVAVMYLMPVN 480
DB 421 GEIVNLSVDAQRFMDLATYINMIMSAPLOVILALYLLMLNLAGSVLAGVAVMYLMPVN 480
QY 481 AVMAKTKTYQVAHMKSKDNRIKLMSEILNGIKVLKYAMELAFKDYALAIROSELKVLK 540
DB 481 AVMAKTKTYQVAHMKSKDNRIKLMSEILNGIKVLKYAMELAFKDYALAIROSELKVLK 540
QY 541 KSAVLSAVGFTTWCCTPFVALCTFAVYVITDENNIIDAQTAVSLAFNLRPPLNLP 600
DB 541 KSAVLSAVGFTTWCCTPFVALCTFAVYVITDENNIIDAQTAVSLAFNLRPPLNLP 600
QY 601 MVISIVQASVSLKRLRIFLSHEELPDSIERRPVKDGGTNSITVENATTFMARSDPT 660
DB 601 MVISIVQASVSLKRLRIFLSHEELPDSIERRPVKDGGTNSITVENATTFMARSDPT 660
QY 661 LMGITFSIPREGALVAVVQVQCGKSLLSALLAEMDKVEGHVAKGSVAVVPOQAMIOND 720
DB 661 LMGITFSIPREGALVAVVQVQCGKSLLSALLAEMDKVEGHVAKGSVAVVPOQAMIOND 720
QY 721 SLRENILFGCQLEPPYRSVYQACALLPDLEILPSGRTIEGEGVNLSSGQKQKRVSLAR 780
DB 721 SLRENILFGCQLEPPYRSVYQACALLPDLEILPSGRTIEGEGVNLSSGQKQKRVSLAR 780
QY 781 AVYENADIYLPDDPLSAVDHVGKHFENYVGPQGMKKNKRIIVTSMSTYLPQVDVIYV 840
DB 781 AVYENADIYLPDDPLSAVDHVGKHFENYVGPQGMKKNKRIIVTSMSTYLPQVDVIYV 840
QY 841 MSGGKISEMGSYQELIARDGAFAEFLRTYASTEQDAEENGVTGVSQPGKEAKOMENG 900
DB 841 MSGGKISEMGSYQELIARDGAFAEFLRTYASTEQDAEENGVTGVSQPGKEAKOMENG 900
QY 901 LVYDSAGKOLOROLSSSSSYSGDISRHNSYAELOKBAKKEEYWKLMADKATQGYKL 960
DB 901 LVYDSAGKOLOROLSSSSSYSGDISRHNSYAELOKBAKKEEYWKLMADKATQGYKL 960
QY 961 SVYWDYKAIGLFISIFLFCMNHVSALASNWYLSMTDDPIYNGTQETHYKRLSYVG 1020
DB 961 SVYWDYKAIGLFISIFLFCMNHVSALASNWYLSMTDDPIYNGTQETHYKRLSYVG 1020
QY 1021 ALGISOGIAVGYSAVSIIGGILASRCIHLVDLHLSILSPMSFFERTPSGNLVNRFSEL 1080
DB 1021 ALGISOGIAVGYSAVSIIGGILASRCIHLVDLHLSILSPMSFFERTPSGNLVNRFSEL 1080
QY 1081 DTVDSMTPEVIKMFMSLFNTGACIVILATPIAAIITPPLGLYFFVQRFYVASSRQL 1140
DB 1081 DTVDSMTPEVIKMFMSLFNTGACIVILATPIAAIITPPLGLYFFVQRFYVASSRQL 1140
QY 1141 KRLSVSRSPVYSHPNETLGVSVIRAPFEEQERTIHOSDLKVDENOKAYYSIVANRWLA 1200
DB 1141 KRLSVSRSPVYSHPNETLGVSVIRAPFEEQERTIHOSDLKVDENOKAYYSIVANRWLA 1200
QY 1201 VRECEVNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLINMLVMSSEMETNIVA 1260

```

Db 1201 VRLCEVNCIYLFALFAVLSRHSLSAGLVLSVSYSLQVTTYLNLVIRMSSEMETNIVA 1260
Qy 1261 VERLKEYSETEKEAPMOIQEAPPSWPQVRFVFRNCCAYRRDLDPVLRHINVTNGG 1320
Db 1261 VERLKEYSETEKEAPMOIQEAPPSWPQVRFVFRNCCAYRRDLDPVLRHINVTNGG 1320
Qy 1321 EKVGIVRTGAGKSSLTLGLFRINESAGEIIGININAKIGLHDLRFKTIIPQDPVLF 1380
Db 1321 EKVGIVRTGAGKSSLTLGLFRINESAGEIIGININAKIGLHDLRFKTIIPQDPVLF 1380
Qy 1381 SGIIRNMUDPFSQYSDSEWVTSLELAHLKDFVSLPDKLDECAEGGENLSVGQRLVCL 1440
Db 1381 SGIIRNMUDPFSQYSDSEWVTSLELAHLKDFVSLPDKLDECAEGGENLSVGQRLVCL 1440
Qy 1441 ARALLRTKTLVDEATRAVDLETDLIQSTIRTOFEDCTVLTAAHRLNTIMDYTRYVL 1500
Db 1441 ARALLRTKTLVDEATRAVDLETDLIQSTIRTOFEDCTVLTAAHRLNTIMDYTRYVL 1500
Qy 1501 DKGEIOYGA PSDLLQORGLFYSMAKDAGLV 1531
Db 1501 DKGEIOYGA PSDLLQORGLFYSMAKDAGLV 1531

```

```

RESULT 8
US-08-463-092B-2
; Sequence 2, Application US/08463092B
; Patent No. 576680
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PARTER RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,092B
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: Q1546
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853

```

```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-092B-2

Query Match 79.3%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MALRGFCSADGSDPLDMANTNTSNPDFFKCFONTLVVWPCFYLMACEPFYLYLSRH 60
Db 1 MALRGFCSADGSDPLDMANTNTSNPDFFKCFONTLVVWPCFYLMACEPFYLYLSRH 60
Qy 61 DRGYIQMTPLNKTKTALGLFLMIVCMADLFYSFWERSRGIFLAVFLVSPTLGITTLLA 120
Db 61 DRGYIQMTPLNKTKTALGLFLMIVCMADLFYSFWERSRGIFLAVFLVSPTLGITTLLA 120
Qy 121 TFLIQLERRKGVQSSGIMLTFWLVYALCALAIRSKIMTLAKEDAOVDLFRDITFYFFS 180
Db 121 TFLIQLERRKGVQSSGIMLTFWLVYALCALAIRSKIMTLAKEDAOVDLFRDITFYFFS 180
Qy 181 LLLIQVLSCFSDRSPLFSETIHDPNPCPSSASFLSRITFMWITGLIVRGYRPLGSD 240
Db 181 LLLIQVLSCFSDRSPLFSETIHDPNPCPSSASFLSRITFMWITGLIVRGYRPLGSD 240
Qy 241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRQPKVYVSSKDPQPKSSKVDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRQPKVYVSSKDPQPKSSKVDANEVEAL 300
Qy 301 IVSPQKEMNPSLFKVLKTFPGFYFLMSFFFKAIHDLMTSGPQILKLIKFNVDYAPD 360
Db 301 IVSPQKEMNPSLFKVLKTFPGFYFLMSFFFKAIHDLMTSGPQILKLIKFNVDYAPD 360
Qy 361 MOGFYTVLLFVTRACQTLVHOFHICFSGMRKIKAVIGAVYRKALVTNSRKSTV 420
Db 361 MOGFYTVLLFVTRACQTLVHOFHICFSGMRKIKAVIGAVYRKALVTNSRKSTV 420
Qy 421 GEIVNLMSVDQRFMDLATYINMIWSAPLOVILALYLMNLGSPVLAVGVAVMVAIPVN 480
Db 421 GEIVNLMSVDQRFMDLATYINMIWSAPLOVILALYLMNLGSPVLAVGVAVMVAIPVN 480
Qy 481 AVPAKTKTYQVAMMSKDNRIKLMEIINGIYVLKYAMELAFKDKVLAIROBELKVLK 540
Db 481 AVPAKTKTYQVAMMSKDNRIKLMEIINGIYVLKYAMELAFKDKVLAIROBELKVLK 540
Qy 541 KSAVLSAVGFTWVCPELVALCTFAYVYTIIDNNILDAQTAFVSLAFNILLRPLNLLP 600
Db 541 KSAVLSAVGFTWVCPELVALCTFAYVYTIIDNNILDAQTAFVSLAFNILLRPLNLLP 600
Qy 601 MVISSIVQASVSLKRLIFLSHELEPDSIERRPVQKGGTNSITVYNAATFTWARSDDPT 660
Db 601 MVISSIVQASVSLKRLIFLSHELEPDSIERRPVQKGGTNSITVYNAATFTWARSDDPT 660
Qy 661 LMGITTSIPRGALVAVVGVQCGKSSLALSALBEMDKVEGHVAKGSVAVYPOQAWIOND 720
Db 661 LMGITTSIPRGALVAVVGVQCGKSSLALSALBEMDKVEGHVAKGSVAVYPOQAWIOND 720
Qy 721 SLRENLIFGQLEBPYRSVIOACALLPDEILIPSGGRTIEGKGVNLSSGQKQKRSVLAR 780
Db 721 SLRENLIFGQLEBPYRSVIOACALLPDEILIPSGGRTIEGKGVNLSSGQKQKRSVLAR 780
Qy 781 AVYSNADIVLFDPLSLAVDAHVGHKIFENYIGKGMKNKTRILLVTHSMSYLPQVDYIIV 840
Db 781 AVYSNADIVLFDPLSLAVDAHVGHKIFENYIGKGMKNKTRILLVTHSMSYLPQVDYIIV 840
Qy 841 MSGKTSBMSYOBLLARDGAFAPLRTYASTOBODAEANGYTGVSQPKAKKONENGM 900
Db 841 MSGKTSBMSYOBLLARDGAFAPLRTYASTOBODAEANGYTGVSQPKAKKONENGM 900
Qy 901 LVTDASGKQIQRLSSSSSYSGDISRHHSNTAEIQKAEAKKEFTWKLMEADKXQDTGVKL 960

```

```

Db      ||||| 901 LVTDSACKOJLOROLSSSSSSSGDLSRHNSIAELQKAEAKKEETWKLMEADKAOTGVCL 960
Qy      961 SVYDYMKAIGLFIPISTIFLPMCNCNVALSNWMLSLMTDDPIYNGOEHTKYRLSYG 1020
Db      961 SVYDYMKAIGLFIPISTIFLPMCNCNVALSNWMLSLMTDDPIYNGOEHTKYRLSYG 1020
Qy      1021 ALGISOGIAVGYMAVSIIGIILASRCLAVDLHLHSILSPMSFPERPESGMLVNRFSKEL 1080
Db      1021 ALGISOGIAVGYMAVSIIGIILASRCLAVDLHLHSILSPMSFPERPESGMLVNRFSKEL 1080
Qy      1081 DTUDSMIPEVTKMWSLFWNIGACIVLILATPIAIIIPPLGLIYFVQRFYVASSQOL 1140
Db      1081 DTUDSMIPEVTKMWSLFWNIGACIVLILATPIAIIIPPLGLIYFVQRFYVASSQOL 1140
Qy      1141 KRLSVSRSPPYSHFNETLLGVSVIRAFEBQERFIHOSDLKVDENOKAYYSIVANRWLA 1200
Db      1141 KRLSVSRSPPYSHFNETLLGVSVIRAFEBQERFIHOSDLKVDENOKAYYSIVANRWLA 1200
Qy      1201 VRLCEVNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMVLVRMSEMETNIVA 1260
Db      1201 VRLCEVNCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMVLVRMSEMETNIVA 1260
Qy      1261 VERLKEVSETEKAPWQIOETAPPSWPOVGRVFRNCLAYREDLDVLRHINVTINGG 1320
Db      1261 VERLKEVSETEKAPWQIOETAPPSWPOVGRVFRNCLAYREDLDVLRHINVTINGG 1320
Qy      1321 EKVGVGTGAGKSLTLGLFRINSAGEIIIDGINIAKIGHDLRKRITIIIPDPLVLF 1380
Db      1321 EKVGVGTGAGKSLTLGLFRINSAGEIIIDGINIAKIGHDLRKRITIIIPDPLVLF 1380
Qy      1381 SGLSRMNLDPFQYSDSEEVMTSLBLAHLKDFVSALPDLDBHCAEGENLSVGORQLVCL 1440
Db      1381 SGLSRMNLDPFQYSDSEEVMTSLBLAHLKDFVSALPDLDBHCAEGENLSVGORQLVCL 1440
Qy      1441 ARALLRKTIIIVDBATVAVLETTDIIQSTIRTOFECCTVLTANRANTIMOTRYVVL 1500
Db      1441 ARALLRKTIIIVDBATVAVLETTDIIQSTIRTOFECCTVLTANRANTIMOTRYVVL 1500
Qy      1501 DKGEIOEYGAPSDLLQORGLFYMAKDGVLV 1531
Db      1501 DKGEIOEYGAPSDLLQORGLFYMAKDGVLV 1531

```

RESULT 9

US-08-462-109A-2

Sequence 2, Application US/08462109A

Patent No. 5882875

GENERAL INFORMATION:

APPLICANT: Cole, Susan P. C.

APPLICANT: Dealey, Roger G.

TITLE OF INVENTION: METHODS FOR IDENTIFYING

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

STREET: 60 State Street, suite 510

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08462,109A

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/966, 923

FILING DATE: 27-OCT-1992

```

APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: POI-002CP4
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-2

```

Query Match 79.3%; Score 7849; DB 1; Length 1531;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy      1 MALRGFCSADGSDPLMDMNTWNTSNPDFTKCFONTVLVWPCEYLMACPEFFLYLSRH 60
Db      1 MALRGFCSADGSDPLMDMNTWNTSNPDFTKCFONTVLVWPCEYLMACPEFFLYLSRH 60
Qy      61 DRGIQMTPLNKTALGFLIMIVCADLPIFSFEBRSRGIFLAVPLVSPILLGITTLLA 120
Db      61 DRGIQMTPLNKTALGFLIMIVCADLPIFSFEBRSRGIFLAVPLVSPILLGITTLLA 120
Qy      121 TFLQLERRKGVSSGIMLTFMLVALCALAIISKIMTALKEDAOYDLDFDITFYVYS 180
Db      121 TFLQLERRKGVSSGIMLTFMLVALCALAIISKIMTALKEDAOYDLDFDITFYVYS 180
Qy      181 LLLIQVLVSCPSDRSPFSEITIHDPNCPRESSASFLRITFMWITGLIVRGYROPLEGSD 240
Db      181 LLLIQVLVSCPSDRSPFSEITIHDPNCPRESSASFLRITFMWITGLIVRGYROPLEGSD 240
Qy      241 LMSLNKEDTSQVVPVLYKMKKECATRKQPVVYVYSSQDPAPKSSSKVDANEVEVAL 300
Db      241 LMSLNKEDTSQVVPVLYKMKKECATRKQPVVYVYSSQDPAPKSSSKVDANEVEVAL 300
Qy      301 IVKSPQKEMNPSLKVLYKTPGFELMSFFPKAIHDLAMPFGPOLKLLIKFVNDTKAPD 360
Db      301 IVKSPQKEMNPSLKVLYKTPGFELMSFFPKAIHDLAMPFGPOLKLLIKFVNDTKAPD 360
Qy      361 WQGYFYTVLFEVTAQTLVLAHQYFHLICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420
Db      361 WQGYFYTVLFEVTAQTLVLAHQYFHLICFVSGMRIKTAVIGAVYRKALVITNSARKSSTV 420
Qy      421 GEIVNLSVDAQRFMDLATYINMTWSAPLOYIALLYLMLNLSVYLAGVAVMYLMPVYN 480
Db      421 GEIVNLSVDAQRFMDLATYINMTWSAPLOYIALLYLMLNLSVYLAGVAVMYLMPVYN 480
Qy      481 AVMAKTKTQVAVAMKSDNRKIKLMEELNGIKYVLKYAMELAFPDKYLAIROBELKYLK 540
Db      481 AVMAKTKTQVAVAMKSDNRKIKLMEELNGIKYVLKYAMELAFPDKYLAIROBELKYLK 540
Qy      541 KSAVLSAVGFTWCTPFLVALCTFAVYVITDENNIIIDAQTAFAVSLAFNLRPPLNLP 600
Db      541 KSAVLSAVGFTWCTPFLVALCTFAVYVITDENNIIIDAQTAFAVSLAFNLRPPLNLP 600
Qy      601 MVISSIVQASVSLKRLRILFISHELEPDSIRRPVKOGGTNSITVRNATPMARSDEPT 660
Db      601 MVISSIVQASVSLKRLRILFISHELEPDSIRRPVKOGGTNSITVRNATPMARSDEPT 660
Qy      661 LNGITFSTPBGALVAVVGVGCGKSLLSALLAEMDKVGGHVALIKGSAAVYVPOQAWIOND 720
Db      661 LNGITFSTPBGALVAVVGVGCGKSLLSALLAEMDKVGGHVALIKGSAAVYVPOQAWIOND 720

```

QY 721 SLRENILFGCOLLEPPYRSVIAQCALLPDLEILPSGRTEIGKGNVLSGGOKORVELAR 780
 DB 721 SLRENILFGCOLLEPPYRSVIAQCALLPDLEILPSGRTEIGKGNVLSGGOKORVELAR 780
 QY 781 AVYSNADIYLPDDPLSAVDHVGKHFENVIQPKMLKNKTRIIIVTHSMSTYLPQVDYIIV 840
 DB 781 AVYSNADIYLPDDPLSAVDHVGKHFENVIQPKMLKNKTRIIIVTHSMSTYLPQVDYIIV 840
 QY 841 MSGKISEMGSYQGLLARDGAFEEFLTYASTEOBDAEENGTVGSGPGKEAQMONGM 900
 DB 841 MSGKISEMGSYQGLLARDGAFEEFLTYASTEOBDAEENGTVGSGPGKEAQMONGM 900
 QY 901 LVYDSAGKOLQROLOSSSSSSSGDISRRHNSTAELOKAKKEEFYWKMEADKATQGVKL 960
 DB 901 LVYDSAGKOLQROLOSSSSSSSGDISRRHNSTAELOKAKKEEFYWKMEADKATQGVKL 960
 QY 961 SVYWDYKKAIGLFISFLISFLFMCNHNVSALASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
 DB 961 SVYWDYKKAIGLFISFLISFLFMCNHNVSALASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
 QY 1021 ALGISOGIAVPGVMAVSIIGIIASRCLHVDLHLSIRSPMSFEERTPSGNLVNRESEKL 1080
 DB 1021 ALGISOGIAVPGVMAVSIIGIIASRCLHVDLHLSIRSPMSFEERTPSGNLVNRESEKL 1080
 QY 1081 DTVDMSIPEVIKMFMSGLFNVIGACIYILLATPIAIIIPPLGLIYFFVORFYVASSROL 1140
 DB 1081 DTVDMSIPEVIKMFMSGLFNVIGACIYILLATPIAIIIPPLGLIYFFVORFYVASSROL 1140
 QY 1141 KRLSEVSRSVPYSHFNFTLLGVSVIRAFEEQERFIHOSDLKVDENQKAYVPSIVANRMLA 1200
 DB 1141 KRLSEVSRSVPYSHFNFTLLGVSVIRAFEEQERFIHOSDLKVDENQKAYVPSIVANRMLA 1200
 QY 1201 VRLCEVNCNIVLPALFAVVISRHSLSGLVGLSVYSLOVTTYANMLVRMSSEKETNIVA 1260
 DB 1201 VRLCEVNCNIVLPALFAVVISRHSLSGLVGLSVYSLOVTTYANMLVRMSSEKETNIVA 1260
 QY 1261 VERLKEYSETEKEAPMOIOETAPPSWPQVGRVFRNYCLRYREDLPVLRHINVTNGG 1320
 DB 1261 VERLKEYSETEKEAPMOIOETAPPSWPQVGRVFRNYCLRYREDLPVLRHINVTNGG 1320
 QY 1321 EKVGIWRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLDLRFKTIIPQDPVLF 1380
 DB 1321 EKVGIWRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLDLRFKTIIPQDPVLF 1380
 QY 1381 SSGSLRMLDPPSQYSDEEWTSLLELAHKPVSALPKLDHECAGEGENTSVGQRQVL 1440
 DB 1381 SSGSLRMLDPPSQYSDEEWTSLLELAHKPVSALPKLDHECAGEGENTSVGQRQVL 1440
 QY 1441 ARALLRKTILVDEATAVADLETDDLIQSTIRPOFEDCTVLTAAHRLNTIMDYTRIVL 1500
 DB 1441 ARALLRKTILVDEATAVADLETDDLIQSTIRPOFEDCTVLTAAHRLNTIMDYTRIVL 1500
 QY 1501 DKGEIOEYGAUSDLLQORGLFYSMAXDAGLV 1531
 DB 1501 DKGEIOEYGAUSDLLQORGLFYSMAXDAGLV 1531

ZIP: K7L 3N6
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/460,907B
 FILING DATE: 05-JUN-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/029,340
 FILING DATE: 8-MAR-1993
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/407,207
 FILING DATE: 20-MAR-1995
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Steeg, Carol Miernicki
 REGISTRATION NUMBER: 39,539
 REFERENCE/DOCKET NUMBER: 01551
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (613) 545-2342
 TELEFAX: (613) 545-6853
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1531 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-460-907B-2
 Query Match 79.3%; Score 7849; DB 1; Length 1531;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MALRGFCSDGSDPLMDWNTVNTSNPDTKCFQNTVLVWPCEYIMACPFYFLYSRH 60
 DB 1 MALRGFCSDGSDPLMDWNTVNTSNPDTKCFQNTVLVWPCEYIMACPFYFLYSRH 60
 QY 61 DRGYIOWTPIANKTKTALGFLMIVCWADLFYSFWERSRGIFLAPVLYVSTLLGITLLA 120
 DB 61 DRGYIOWTPIANKTKTALGFLMIVCWADLFYSFWERSRGIFLAPVLYVSTLLGITLLA 120
 QY 121 TFLIOLERRKGVSSGIMLTFMVLVCAIALIRSKIMTALKBDAQVDLFRDITFYVFS 180
 DB 121 TFLIOLERRKGVSSGIMLTFMVLVCAIALIRSKIMTALKBDAQVDLFRDITFYVFS 180
 QY 181 LLLIOLVLSGFSRSPFLFSETIHDPNPCPESSASFLRITFWWITGLIVRGYQPLEGSD 240
 DB 181 LLLIOLVLSGFSRSPFLFSETIHDPNPCPESSASFLRITFWWITGLIVRGYQPLEGSD 240
 QY 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKYVYSSQDPQOPRESSRYDANEVEAL 300
 DB 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKYVYSSQDPQOPRESSRYDANEVEAL 300
 QY 301 IVSPQKEMNPSLFKVLVYKTFPGYFLMSFFFAIHDLMFSGQIILKLKIFVNDTKAPD 360
 DB 301 IVSPQKEMNPSLFKVLVYKTFPGYFLMSFFFAIHDLMFSGQIILKLKIFVNDTKAPD 360
 QY 361 WGGYFTVLLFVACIQTIVLHQYFHI CFVSGRIRITAVIGAVYRRAVLTNSARSSSTV 420
 DB 361 WGGYFTVLLFVACIQTIVLHQYFHI CFVSGRIRITAVIGAVYRRAVLTNSARSSSTV 420

RESULT 10
 US-08-460-907B-2
 Sequence 2, Application US/08460907B
 Patent No. 5891724
 GENERAL INFORMATION:
 APPLICANT: Deeley, Roger G.
 APPLICANT: Cole, Susan P.C.
 TITLE OF INVENTION: METHODS FOR CONFERRING MCL1DRUG
 TITLE OF INVENTION: RESISTANCE ON A CELL
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: PARTEO RESEARCH & DEVELOPMENT INNOVATIONS
 STREET: Queen's University at Kingston
 CITY: Kingston
 STATE: Ontario
 COUNTRY: CANADA

```

QY 421 GEIYNLMSVDAQRPMDLATYINMTWSAPLOVTLALYILMLNGPSVLGAVAMVLMVAVN 480
DB 421 GEIYNLMSVDAQRPMDLATYINMTWSAPLOVTLALYILMLNGPSVLGAVAMVLMVAVN 480
QY 481 AVAMAKTKTYOVAMKSKDNRIKIMNELINGIKYLKLYAMELAKDKVLAIROBELKVLK 540
DB 481 AVAMAKTKTYOVAMKSKDNRIKIMNELINGIKYLKLYAMELAKDKVLAIROBELKVLK 540
QY 541 KSAVLSAVGFTWCTPFLVALCTFAVYVITDENNILLDACTAFVSLAFNLRPELNLIP 600
DB 541 KSAVLSAVGFTWCTPFLVALCTFAVYVITDENNILLDACTAFVSLAFNLRPELNLIP 600
QY 601 MTSSIVQASVSLKRLRIFLSHEELBPDSIRRPVKOGGNTSTVRNATTTWARSDDPT 660
DB 601 MTSSIVQASVSLKRLRIFLSHEELBPDSIRRPVKOGGNTSTVRNATTTWARSDDPT 660
QY 661 LINGITFSPGALVAVVGVGCGKSLSLALLAEMDKVGHVALKGSAAVYPOQAMQND 720
DB 661 LINGITFSPGALVAVVGVGCGKSLSLALLAEMDKVGHVALKGSAAVYPOQAMQND 720
QY 721 SLRENIIFGCOLPEPYRSVIOACALLPDLILPSGDRTEIGEGVNLSGGOKORVSLAR 780
DB 721 SLRENIIFGCOLPEPYRSVIOACALLPDLILPSGDRTEIGEGVNLSGGOKORVSLAR 780
QY 781 AVYSNADIYLFDDPLSAVDAHVKHIFENVIGPKGMLKNKTRIIIVTSHMSYLPQVDVIV 840
DB 781 AVYSNADIYLFDDPLSAVDAHVKHIFENVIGPKGMLKNKTRIIIVTSHMSYLPQVDVIV 840
QY 841 MSGGKISMSGYOBLIARDGAFARLRTTYASTEOBDAEENGVCVSGPGEAKOMENGM 900
DB 841 MSGGKISMSGYOBLIARDGAFARLRTTYASTEOBDAEENGVCVSGPGEAKOMENGM 900
QY 901 LVTDSAGKOLOROLSSSSSSSGDISRHHNSTAELOKAKKEETKLMLEADKAOQGVKL 960
DB 901 LVTDSAGKOLOROLSSSSSSSGDISRHHNSTAELOKAKKEETKLMLEADKAOQGVKL 960
QY 961 SVYWDYMAKIGLFLSFLSIFLPMCNHVSALASNTWLSMTDPIVNGTQHTKVLASYG 1020
DB 961 SVYWDYMAKIGLFLSFLSIFLPMCNHVSALASNTWLSMTDPIVNGTQHTKVLASYG 1020
QY 1021 ALGISOGIAVRCGYMAVSIIGIILASRCLHVDLHLSILASPMSPFERTSGMLVNRFSSEL 1080
DB 1021 ALGISOGIAVRCGYMAVSIIGIILASRCLHVDLHLSILASPMSPFERTSGMLVNRFSSEL 1080
QY 1081 DTYSMTIPEVIMKMGSLFNVIYACIVILLATPIAIIIPGLIYFFVQRFYVASSROL 1140
DB 1081 DTYSMTIPEVIMKMGSLFNVIYACIVILLATPIAIIIPGLIYFFVQRFYVASSROL 1140
QY 1141 KRLSVSRSPVYSHFNETLLGVSIVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLSVSRSPVYSHFNETLLGVSIVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLFCVGCIVLFAALFAVISRHSLSAGLVGSVYSIQVTTYLNMVLRMSSEMETNIVA 1260
DB 1201 VRLFCVGCIVLFAALFAVISRHSLSAGLVGSVYSIQVTTYLNMVLRMSSEMETNIVA 1260
QY 1261 VERLKEVETEKEAPWMOIETRPSPSPQVGRVERRANCRLKRELDVLDKHNINVTINGG 1320
DB 1261 VERLKEVETEKEAPWMOIETRPSPSPQVGRVERRANCRLKRELDVLDKHNINVTINGG 1320
QY 1321 EKVGI VGTGAKSSLTGLFRINSABGEI11IDGINIAKIGLHDLRFKRTI11PODPVLF 1380
DB 1321 EKVGI VGTGAKSSLTGLFRINSABGEI11IDGINIAKIGLHDLRFKRTI11PODPVLF 1380
QY 1381 SSGSLRMLNDPSPQSVDEEVTSLSLAHLKDFVSALPDKLDHCCAGGENTLSVGQOLVCL 1440
DB 1381 SSGSLRMLNDPSPQSVDEEVTSLSLAHLKDFVSALPDKLDHCCAGGENTLSVGQOLVCL 1440
QY 1441 ARALLRKTIIYLDATAVAVDLETDLLIOSTIRQFEDCTVLTIAHRLNTIMDYRIVVL 1500
DB 1441 ARALLRKTIIYLDATAVAVDLETDLLIOSTIRQFEDCTVLTIAHRLNTIMDYRIVVL 1500
QY 1501 DKGEIQEYGAPSDLLQGRGLFYSMAKDAGLV 1531

```

```

DB 1501 DKGEIQEYGAPSDLLQGRGLFYSMAKDAGLV 1531

RESULT 11
US-08-463-179A-2
; Sequence 2, Application US/08463179A
; Patent No. 6001563
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,179A
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Decont, Giulio A. Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: P01-002CP8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-179A-2

Query Match 79.3%; Score 7849; DB 2; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALGFCSDGSDLLMDNMVNTSNPDTKCFONTVLVWVPCFYLMACPFYFLYSRH 60
DB 1 MALGFCSDGSDLLMDNMVNTSNPDTKCFONTVLVWVPCFYLMACPFYFLYSRH 60
QY 61 DRGYIOMTPLUNKKTALGFLIMIVCMADLFYSFWERSRGIFLAEVFLVSPTLIGITTLA 120
DB 61 DRGYIOMTPLUNKKTALGFLIMIVCMADLFYSFWERSRGIFLAEVFLVSPTLIGITTLA 120
QY 121 TFLIQLERRKVGSSGIMLTFWLVALYCALAITSKIMTALKEDAQVDLFPDITFYVYFS 180
DB 121 TFLIQLERRKVGSSGIMLTFWLVALYCALAITSKIMTALKEDAQVDLFPDITFYVYFS 180
QY 181 LLLIQLVLSGSDSPFSEITIHDPNCPSSASFLSITFWMTTGLIVRGYROPLEBSD 240
DB 181 LLLIQLVLSGSDSPFSEITIHDPNCPSSASFLSITFWMTTGLIVRGYROPLEBSD 240

```

QY 241 LMSLNKEDTSEQVVPVLVKMKKECAKTRKQPVKVTYSSKDPAPQPKSSKVDANEVEAL 300
 DB 241 LMSLNKEDTSEQVVPVLVKMKKECAKTRKQPVKVTYSSKDPAPQPKSSKVDANEVEAL 300
 QY 301 IVKSPQKEMNPSPLEKVLKTFGPFYFLMSFPFKAIHDLMPFSGPOLIKLIFVNDTRAPD 360
 DB 301 IVKSPQKEMNPSPLEKVLKTFGPFYFLMSFPFKAIHDLMPFSGPOLIKLIFVNDTRAPD 360
 QY 361 WQGFYVVLVFNVAQCTVLVHOFPHICFVSGNMKIKNAVIGAVYRKALVITNSARKSTV 420
 DB 361 WQGFYVVLVFNVAQCTVLVHOFPHICFVSGNMKIKNAVIGAVYRKALVITNSARKSTV 420
 QY 421 GEIYNLMSVDAOREFMDLATYINMWSAPLOYITALLYLMLNGPSVLAVAVWVLAMPVN 480
 DB 421 GEIYNLMSVDAOREFMDLATYINMWSAPLOYITALLYLMLNGPSVLAVAVWVLAMPVN 480
 QY 481 AVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYLAMELAFKDYLAIROBELKVLK 540
 DB 481 AVAMKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYLAMELAFKDYLAIROBELKVLK 540
 QY 541 KSAVLAVGTFTWCTPPLVALCTFPAYVYITDENNIIDAQTAFLVALFNILRPLNLP 600
 DB 541 KSAVLAVGTFTWCTPPLVALCTFPAYVYITDENNIIDAQTAFLVALFNILRPLNLP 600
 QY 601 MWISSIVQASVSLKRLIFLSHELEPDSIERRPVKDGGGNSITVANAFTTARSDBPT 660
 DB 601 MWISSIVQASVSLKRLIFLSHELEPDSIERRPVKDGGGNSITVANAFTTARSDBPT 660
 QY 661 LMGITFSPREGALVAVVGVGCGKSLLSALLAEMDVKEGHVAKGSVAVYVPOQAWIOND 720
 DB 661 LMGITFSPREGALVAVVGVGCGKSLLSALLAEMDVKEGHVAKGSVAVYVPOQAWIOND 720
 QY 721 SLRBNILFGCQLEPPYRSYIQAACALLPDLFILPSGRTEIGEKNVLSGGOKRVSLAR 780
 DB 721 SLRBNILFGCQLEPPYRSYIQAACALLPDLFILPSGRTEIGEKNVLSGGOKRVSLAR 780
 QY 781 AVYSNADIYLPDPLSAVDHVGHIENYIGPMKMKNTRIILVTHSMVSLPVDYIIV 840
 DB 781 AVYSNADIYLPDPLSAVDHVGHIENYIGPMKMKNTRIILVTHSMVSLPVDYIIV 840
 QY 841 MSGGKISEMGSYOELARDGAFELRTYASTEOBOAEENGVTGSGPGEAKOMENGM 900
 DB 841 MSGGKISEMGSYOELARDGAFELRTYASTEOBOAEENGVTGSGPGEAKOMENGM 900
 QY 901 LVYDSAGKQIQRULSSSSSYSGDISRHNSHTAELOKAEAKKEETKLMLEADKAQTQVKL 960
 DB 901 LVYDSAGKQIQRULSSSSSYSGDISRHNSHTAELOKAEAKKEETKLMLEADKAQTQVKL 960
 QY 961 SVYWDVWKATGLFISFLIFLPMCNHVSALASNWLSLMTDDPIVNGTOEHTKRLSVYG 1020
 DB 961 SVYWDVWKATGLFISFLIFLPMCNHVSALASNWLSLMTDDPIVNGTOEHTKRLSVYG 1020
 QY 1021 ALGISOGIAVFGYSMAVISGIIASRCIAYDLHSILRSPMSPERPSPGNLVNRFSEKL 1080
 DB 1021 ALGISOGIAVFGYSMAVISGIIASRCIAYDLHSILRSPMSPERPSPGNLVNRFSEKL 1080
 QY 1081 DTVDMSMIPBVIKPMFMSLFNVIGACIYILLATPIAIIIPPLGIYFVORFYVASSROL 1140
 DB 1081 DTVDMSMIPBVIKPMFMSLFNVIGACIYILLATPIAIIIPPLGIYFVORFYVASSROL 1140
 QY 1141 KRLSVRSRSPVSHFNFTLLGVSVIRAFEEQERFIHOSDKVDENQAKYPSIYANWLA 1200
 DB 1141 KRLSVRSRSPVSHFNFTLLGVSVIRAFEEQERFIHOSDKVDENQAKYPSIYANWLA 1200
 QY 1201 VRLCEVNCIYVFAALFAVISRHSLSAGLVGLSVSYLSQVTTYANMIVRMSSEMETNIVA 1260
 DB 1201 VRLCEVNCIYVFAALFAVISRHSLSAGLVGLSVSYLSQVTTYANMIVRMSSEMETNIVA 1260
 QY 1261 VERLKEYSSETEKAPWQIOETAPSSWPQVGRVAFENYCLRYREDDLVFLRHINVTINGG 1320
 DB 1261 VERLKEYSSETEKAPWQIOETAPSSWPQVGRVAFENYCLRYREDDLVFLRHINVTINGG 1320
 QY 1321 EKVGIVORTAGKGSLSLGLFRINESABGEIIDGINIAKIGLHDLAFKTIITIPQDVLF 1380

DB 1321 EKVGIVORTAGKGSLSLGLFRINESABGEIIDGINIAKIGLHDLAFKTIITIPQDVLF 1380
 QY 1381 SGSLRMLNLPDPSQYDSEEVTSLELAHLKDFVSLPKDLHBCAEGGENTSVGQRLVCL 1440
 DB 1381 SGSLRMLNLPDPSQYDSEEVTSLELAHLKDFVSLPKDLHBCAEGGENTSVGQRLVCL 1440
 QY 1441 ARALLRRTKTLVDEATAVADLETFDILIOSTIRTOPEBDCTVLTARHLNTIMOTRYVL 1500
 DB 1441 ARALLRRTKTLVDEATAVADLETFDILIOSTIRTOPEBDCTVLTARHLNTIMOTRYVL 1500
 QY 1501 DKGEIOEYAPSDILOQRGLFYSGMAXDAGLV 1531
 DB 1501 DKGEIOEYAPSDILOQRGLFYSGMAXDAGLV 1531
 RESULT 12
 US-08-461-384B-2
 : Sequence 2, Application US/08461384B
 : Patent No. 6025473
 : GENERAL INFORMATION:
 : APPLICANT: Cole, Susan P.C.
 : APPLICANT: Deeley, Roger G.
 : TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
 : NUMBER OF SEQUENCES: 10
 : CORRESPONDENCE ADDRESS:
 : ADDRESS: PARTO RESEARCH & DEVELOPMENT INNOVATIONS
 : STREET: Queen's University at Kingston
 : CITY: Kingston
 : STATE: Ontario
 : COUNTRY: CANADA
 : ZIP: K7L 3N6
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: ASCII text
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/461,384B
 : FILING DATE: 05-JUN-95
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 07/966,923
 : FILING DATE: 27-OCT-1992
 : APPLICATION NUMBER: 08/029,340
 : FILING DATE: 8-MAR-1993
 : APPLICATION NUMBER: 08/141,893
 : FILING DATE: 26-OCT-1993
 : APPLICATION NUMBER: 08/407,207
 : FILING DATE: 20-MAR-1995
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Steeg, Carol Miernicki
 : REGISTRATION NUMBER: 39,539
 : REFERENCE/DOCKET NUMBER: 01547
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (613) 545-2342
 : TELEFAX: (613) 545-6853
 : INFORMATION FOR SEQ ID NO: 2:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1531 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : US-08-461-384B-2
 Query Match 79.3%; Score 7849; DB 2; Length 1531;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MALRGFCSDAGSDPLMDMNTVNTSNPDFTKCFONTLVVWPCFYIMACPPFYFLYSRH 60
 DB 1 MALRGFCSDAGSDPLMDMNTVNTSNPDFTKCFONTLVVWPCFYIMACPPFYFLYSRH 60
 QY 61 DRGYIGMTPLNKTKTLGLFLMIVCNADLFYFWRERSGIFLAPVFLVSPFTLLGITTLLA 120

Db 61 DRGTYQMTPLNKTKTALGFLIMICVMDLFFYSFMRSGKIFLAVFVLSPTLIGITLLA 120
Qy 121 TPLIQERRRGVSSGIMLTFMVALVCAALILRSKIMTALKEADVDLDFEDIFYYFS 180
Db 121 TPLIQERRRGVSSGIMLTFMVALVCAALILRSKIMTALKEADVDLDFEDIFYYFS 180
Qy 181 LLLIQVLSCFSDRSPLFSETHDNPDPSSASFLSHITFMTITGLIVRGYRPLBGS 240
Db 181 LLLIQVLSCFSDRSPLFSETHDNPDPSSASFLSHITFMTITGLIVRGYRPLBGS 240
Qy 241 LWSLANKEDTSRQVVPVLYKWKKECAKTRKOPVYVYSSKDPAPKSSSKVDANEVEAL 300
Db 241 LWSLANKEDTSRQVVPVLYKWKKECAKTRKOPVYVYSSKDPAPKSSSKVDANEVEAL 300
Qy 301 IVSPKEMNPSLKFVLYKTEGPFILMSFFKALHDLMFSGPOLIKLLIFVNDTAPD 360
Db 301 IVSPKEMNPSLKFVLYKTEGPFILMSFFKALHDLMFSGPOLIKLLIFVNDTAPD 360
Qy 361 WQGYFYVLLFVTACLQTLVLHQYFHLCFVSGMRITKAVIGAVYRKALVITNSARKSTV 420
Db 361 WQGYFYVLLFVTACLQTLVLHQYFHLCFVSGMRITKAVIGAVYRKALVITNSARKSTV 420
Qy 421 GEIYNLMSVDAQRPMDLATYINMTWSAPLOYITALLYLMLNGPSVLAGAVVMTLMPVN 480
Db 421 GEIYNLMSVDAQRPMDLATYINMTWSAPLOYITALLYLMLNGPSVLAGAVVMTLMPVN 480
Qy 481 AVNMAKTKTYOVAMHKSNDRIKLMNEILNGIKVLYAMELAFDKYLAIRQELKYLK 540
Db 481 AVNMAKTKTYOVAMHKSNDRIKLMNEILNGIKVLYAMELAFDKYLAIRQELKYLK 540
Qy 541 KSAVLASVGTFTWCTPFLVALCTFAVYVITDENNILDQAFAVSLFNILRPPLNLP 600
Db 541 KSAVLASVGTFTWCTPFLVALCTFAVYVITDENNILDQAFAVSLFNILRPPLNLP 600
Qy 601 MVTSSIVQASVSLKRLRIFLSHEBELPDSIRRPYKOGGNSITVRNATTTWARSDEPT 660
Db 601 MVTSSIVQASVSLKRLRIFLSHEBELPDSIRRPYKOGGNSITVRNATTTWARSDEPT 660
Qy 661 LINGITFSLPEGALVAVVQVCGKSLLSALLAENDKVEGHVALKGSAAVYPOQAMIOND 720
Db 661 LINGITFSLPEGALVAVVQVCGKSLLSALLAENDKVEGHVALKGSAAVYPOQAMIOND 720
Qy 721 SLRENILFGCOLLEPPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSSGQOKRVSLAR 780
Db 721 SLRENILFGCOLLEPPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLSSGQOKRVSLAR 780
Qy 781 AVYSNADIYLFDDPLSAVDAHVGHIFENVIQPKMLKNKTRILVTHSMSTLPQVDVITV 840
Db 781 AVYSNADIYLFDDPLSAVDAHVGHIFENVIQPKMLKNKTRILVTHSMSTLPQVDVITV 840
Qy 841 MSGGKISMSGYOELLARDGAFABELRTYASTEOBDAEENGUVGVSPPGKEAKOMENGM 900
Db 841 MSGGKISMSGYOELLARDGAFABELRTYASTEOBDAEENGUVGVSPPGKEAKOMENGM 900
Qy 901 LVTPDSAGKOLOROLSSSSSSSGDISRHHNSTAELOKAKAKEETKLMLEADKAOQGYKL 960
Db 901 LVTPDSAGKOLOROLSSSSSSSGDISRHHNSTAELOKAKAKEETKLMLEADKAOQGYKL 960
Qy 961 SVYWDYMKAIQLFISFLSIFLPMCNHVSALASNTWLSMTDDPIVNGTQEHTRVLSYVG 1020
Db 961 SVYWDYMKAIQLFISFLSIFLPMCNHVSALASNTWLSMTDDPIVNGTQEHTRVLSYVG 1020
Qy 1021 ALGISOGIAVFGYMAVSIIGIILASRCHVDLHLSILSPMSFFERTSSGULVNFSEKL 1080
Db 1021 ALGISOGIAVFGYMAVSIIGIILASRCHVDLHLSILSPMSFFERTSSGULVNFSEKL 1080
Qy 1081 DTVDSMIPEVIMKMGSLFNVYIACITVILLATPIAIIIPPLGLYFFVQRPYVASRQOL 1140
Db 1081 DTVDSMIPEVIMKMGSLFNVYIACITVILLATPIAIIIPPLGLYFFVQRPYVASRQOL 1140
Qy 1141 KRLESVSRSPYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
Db 1141 KRLESVSRSPYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200

Db 1141 KRLESVSRSPYSHNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
Qy 1201 VRLECVGNICIVLFAALFAVISRHSLSAGLVGSYSLOVTTYLNLMLVRMSSEMETNIVA 1260
Db 1201 VRLECVGNICIVLFAALFAVISRHSLSAGLVGSYSLOVTTYLNLMLVRMSSEMETNIVA 1260
Qy 1261 VERLKEYSETEKEAPWIOETAPSSWPQYGRVFRNYCIRYREDDLPVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKEAPWIOETAPSSWPQYGRVFRNYCIRYREDDLPVLRHINVTINGG 1320
Qy 1321 EKVGIVERTGAGKSLTLGLFRINESAGEITIIDGINIAKIGLHDLAFKTIITIPODPLF 1380
Db 1321 EKVGIVERTGAGKSLTLGLFRINESAGEITIIDGINIAKIGLHDLAFKTIITIPODPLF 1380
Qy 1381 SGLSRMMLDPFSQYSDDEWVTSLELAHKDFVSALPKLDHECAGGENLSVGOQOLVCL 1440
Db 1381 SGLSRMMLDPFSQYSDDEWVTSLELAHKDFVSALPKLDHECAGGENLSVGOQOLVCL 1440
Qy 1441 ARALLRRTKILVIDEATAVADLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Db 1441 ARALLRRTKILVIDEATAVADLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500
Qy 1501 DKGEIOEYGA PSDILOQRGLFYSMAKDAGLV 1531
Db 1501 DKGEIOEYGA PSDILOQRGLFYSMAKDAGLV 1531

RESULT 13
US-08-407-207A-2
Sequence 2, Application US/08407207A
Patent No. 6063621
GENERAL INFORMATION:
Applicant: Deeley, Roger G.
Applicant: Cole, Susan P.C.
TITLE OF INVENTION: ANTIBODIES TO A MULTIDRUG RESISTANCE PROTEIN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESSES:
ADDRESSEE: PARTER RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/407,207A
FILING DATE: 20-MAR-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01512
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-407-207A-2

Query Match 79.3%; Score 7849; Db 2; Length 1531;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy 1 MALRGFCGADSGDPLMDNMTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFYLYLSRH 60
Db 1 MALRGFCGADSGDPLMDNMTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFYLYLSRH 60
Qy 61 DRGYQMTPLNKTATLGLFLMTIYCMADLFYSFWRSGIPLAPFLVSPITLIGTTLLA 120
Db 61 DRGYQMTPLNKTATLGLFLMTIYCMADLFYSFWRSGIPLAPFLVSPITLIGTTLLA 120
Qy 121 TFLIOLERRKGVSSGIMLTFMLVALVCALATIRSKITATLKEDAQVDLFRDITYYVES 180
Db 121 TFLIOLERRKGVSSGIMLTFMLVALVCALATIRSKITATLKEDAQVDLFRDITYYVES 180
Qy 181 LLLIQLVLSCFSDRSPLFSETHDBNCPESASASLSRIITFWITGLIVRGYRQPLEGSD 240
Db 181 LLLIQLVLSCFSDRSPLFSETHDBNCPESASASLSRIITFWITGLIVRGYRQPLEGSD 240
Qy 241 LMSLNKETSQVVPVLVKNWKKCEAKTRKQPVKYVSSKQPAQKSSKYDANEVEAL 300
Db 241 LMSLNKETSQVVPVLVKNWKKCEAKTRKQPVKYVSSKQPAQKSSKYDANEVEAL 300
Qy 301 IVKSPQKEMNPSLFVLYKTFGPYFLMSPFFKAIHDLMMPSGPOILKLIKPVNDTKAD 360
Db 301 IVKSPQKEMNPSLFVLYKTFGPYFLMSPFFKAIHDLMMPSGPOILKLIKPVNDTKAD 360
Qy 361 WQGYFYTLFVTAQTLQTLVLAHQYFHLICFVSGMRKTAIVAGVYRKALVITNSARKSTV 420
Db 361 WQGYFYTLFVTAQTLQTLVLAHQYFHLICFVSGMRKTAIVAGVYRKALVITNSARKSTV 420
Qy 421 GEIYNLMSVDAQRFMDLATYINMTISAPLQVITLALYILMLNGSVLAGVAVMLMVEVN 480
Db 421 GEIYNLMSVDAQRFMDLATYINMTISAPLQVITLALYILMLNGSVLAGVAVMLMVEVN 480
Qy 481 AVAMKRTKYOVAMHMSKONRIKLMNELNGIKYKUYAMELAPKDKTALROBELKYLK 540
Db 481 AVAMKRTKYOVAMHMSKONRIKLMNELNGIKYKUYAMELAPKDKTALROBELKYLK 540
Qy 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNIIDAQTAFAVSLAFNLIRFPLNLP 600
Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNIIDAQTAFAVSLAFNLIRFPLNLP 600
Qy 601 MWISSIYOASVSLKRLRIFLSHBELEPDSIRRPYKGGGJNSITVRNATFTWASDPPT 660
Db 601 MWISSIYOASVSLKRLRIFLSHBELEPDSIRRPYKGGGJNSITVRNATFTWASDPPT 660
Qy 661 LINGITFSPBGAIVAVVGVCGGKSLSLALLAEMDKVEGVATKGSVAAYVPOQAMTND 720
Db 661 LINGITFSPBGAIVAVVGVCGGKSLSLALLAEMDKVEGVATKGSVAAYVPOQAMTND 720
Qy 721 SLRENILFGCOLLEPPYRSVIOACALLPDLILPSGDRTEIGEGVNLSSGOKORVSLAR 780
Db 721 SLRENILFGCOLLEPPYRSVIOACALLPDLILPSGDRTEIGEGVNLSSGOKORVSLAR 780
Qy 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIQPKMLKNKTRILVTHSMSTLPQVDVILV 840
Db 781 AVYSNADIYLFDDPLSAVDHVGKHFENVIQPKMLKNKTRILVTHSMSTLPQVDVILV 840
Qy 841 MSGKISGMSYQELLARDGAFAPFLRTYASTBOBOAENGCVGSGPGEAQAOMENGM 900
Db 841 MSGKISGMSYQELLARDGAFAPFLRTYASTBOBOAENGCVGSGPGEAQAOMENGM 900
Qy 901 LVYDSAGKOLROULSSSSSYSGDISRHNSTAELQKAEKKEETWKLMEADKAOTGOVKL 960
Db 901 LVYDSAGKOLROULSSSSSYSGDISRHNSTAELQKAEKKEETWKLMEADKAOTGOVKL 960
Qy 961 SYVMDYMKALGLFISFLSIFLPCNHNVSALASNTWLSMTDDPIVNGTQEHTKYRLSYG 1020
Db 961 SYVMDYMKALGLFISFLSIFLPCNHNVSALASNTWLSMTDDPIVNGTQEHTKYRLSYG 1020

```

```

Qy 1021 ALGISOGIAVFGYMAVSIIGILASRCILVLDLHLSILSPMSFFERTPSGNLVNRFSEKL 1080
Db 1021 ALGISOGIAVFGYMAVSIIGILASRCILVLDLHLSILSPMSFFERTPSGNLVNRFSEKL 1080
Qy 1081 DTVDSMIPEVIKMFMSLNFNIGACIVILLATPIAIIIPPLGLIYFFVQRFYVASSRQL 1140
Db 1081 DTVDSMIPEVIKMFMSLNFNIGACIVILLATPIAIIIPPLGLIYFFVQRFYVASSRQL 1140
Qy 1141 KRLSVSRSPYSHFNELTGVSVIRAFEBERTRIHOGLDKVDENOKRAYPSIVANRWLA 1200
Db 1141 KRLSVSRSPYSHFNELTGVSVIRAFEBERTRIHOGLDKVDENOKRAYPSIVANRWLA 1200
Qy 1201 VRLCEVGNCLVFPALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRSSEMETNIVA 1260
Db 1201 VRLCEVGNCLVFPALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRSSEMETNIVA 1260
Qy 1261 VERLKEYSETEKAPMOIOTETAPSSPQVGRVBFRRNYCLYRREDLDFVLRIHINTINGG 1320
Db 1261 VERLKEYSETEKAPMOIOTETAPSSPQVGRVBFRRNYCLYRREDLDFVLRIHINTINGG 1320
Qy 1321 EKVGIVGRTGAGKSLTLGLFRINESABGEIINDGINIAKIGLHDLRFKTIIPQDPVLF 1380
Db 1321 EKVGIVGRTGAGKSLTLGLFRINESABGEIINDGINIAKIGLHDLRFKTIIPQDPVLF 1380
Qy 1381 SGSLRMLNDPFSQYSDSEVWTSLELAHLKDPVSALPKLDHECAEGGENLSVGOROLVCL 1440
Db 1381 SGSLRMLNDPFSQYSDSEVWTSLELAHLKDPVSALPKLDHECAEGGENLSVGOROLVCL 1440
Qy 1441 ABALLRKTILVDEATPAVDLETDDLIOSTIRTOFEDCTVLTAAHRLNTIMDYTRVIVL 1500
Db 1441 ABALLRKTILVDEATPAVDLETDDLIOSTIRTOFEDCTVLTAAHRLNTIMDYTRVIVL 1500
Qy 1501 DKGEIOEYGA PSDLLQORGLFYMAKDAGLV 1531
Db 1501 DKGEIOEYGA PSDLLQORGLFYMAKDAGLV 1531

```

```

RESULT 14
US-08-463-092B-6
; Sequence 6, Application US/08463092B
; Patent No. 5766880
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P. C.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,092B
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/141,893

```

FILING DATE: 26-OCT-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/407,207
 FILING DATE: 20-MAR-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Steeg, Carol Miernicki
 REGISTRATION NUMBER: 39,539
 REFERENCE/DOCKET NUMBER: Q1546
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (613) 545-2342
 TELEFAX: (613) 545-6853
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1528 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-463-092B-6

Query Match 70.7%; Score 7002.5; DB 1; Length 1528;
 Best Local Similarity 88.0%; Pred. No. 0;
 Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRGFCGADGSDPLMDNVTMTNSPDFTKCFQNTVLVWVPCFYLMACPEPFYLYLSRH 60
 DB 1 MALRSCGADGSDPLMDNVTMTNSPDFTKCFQNTVLVWVPCFYLMACPEPFYLYLSRH 60
 QY 61 DRGYQMTPLNKTALGFLLMTVCMADLFYSFMRSGIGFLAPFLVSPITLITL 120
 DB 61 DRGYQMTPLNKTALGFLLMTVCMADLFYSFMRSGIGFLAPFLVSPITLITL 120
 QY 121 TFLQLBERKVOSSGIMLTFMLVALVICALILRSKITALKEDAVDLFRDITFYYS 180
 DB 121 TFLQLBERKVOSSGIMLTFMLVALVICALILRSKITALKEDAVDLFRDITFYYS 180
 QY 181 LLLIQLVLSCFPSDRSPLFSETHIDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240
 DB 181 LVLVOLVLSCFPSDRSPLFSETHIDNPCESSASFLSRITFWITGLIVRGYRQPLEGSD 240
 QY 241 LMSLNKETSQVNPVLYVKNKKKCAKTRKQPVKYVY- SKDPAQPKSSSKVDANEVEA 299
 DB 241 LMSLNKETSQVNPVLYVKNKKKCAKTRKQPVKYVY- SKDPAQPKSSSKVDANEVEA 299
 QY 300 LIVSPKEMWPSLPEKVLTKTFGPFELMSFFPKAHDLMFSGPOILKILIKFVNDTAP 359
 DB 301 LIVSPKEMWPSLPEKVLTKTFGPFELMSFFPKAHDLMFSGPOILKILIKFVNDTAP 359
 QY 360 DMQGYFTVLLFVTAQTLVLAHQYFHCVPSGMRKTAIVIGAVYRKALVITNSARKSST 419
 DB 361 DMQGYFTVLLFVTAQTLVLAHQYFHCVPSGMRKTAIVIGAVYRKALVITNSARKSST 419
 QY 420 VGEIYNLMSVDAQRPMDLATTYINMTWSAPLOYITALLYLMLNGSVTLAGVAVMLAMPV 479
 DB 421 VGEIYNLMSVDAQRPMDLATTYINMTWSAPLOYITALLYLMLNGSVTLAGVAVMLAMPV 479
 QY 480 NAVNANKTKTYOVAMKSKDNRIKIMNELLINGIKYLKLYAMELAKDKVLAIRSEELVYL 539
 DB 481 NAVNANKTKTYOVAMKSKDNRIKIMNELLINGIKYLKLYAMELAKDKVLAIRSEELVYL 539
 QY 540 KKSAYLSAVGTFTWCTPFVLAICTFAVVTIDENNIIIDAQTAFAFSLAFNIRFPLNLT 599
 DB 541 KKSAYLSAVGTFTWCTPFVLAICTFAVVTIDENNIIIDAQTAFAFSLAFNIRFPLNLT 599
 QY 600 PMVTSISYQASVSLKRLRIFLSHELEPDSIRRRPVKGGGNSITVNNATFTWASDP 659
 DB 601 PMVTSISYQASVSLKRLRIFLSHELEPDSIRRRPVKGGGNSITVNNATFTWASDP 659
 QY 660 TLNGTTFSPREGALVAVVGVCGGKSLSLALAMDVEGVHVAIKGSVAVYPOQAMION 719
 DB 660 TLNGTTFSPREGALVAVVGVCGGKSLSLALAMDVEGVHVAIKGSVAVYPOQAMION 719

QY 720 DSLRENTLFGCQLEPPYRYTQACALLPDLEILLPSGDRTEIGKGVNLSGGQQRVSLA 779
 DB 720 DSLRENTLFGHPLQENNYKAMKAMECALLPDLEILLPSGDRTEIGKGVNLSGGQQRVSLA 779
 QY 780 RAVYSNADIVLFDPLSNVDHAKKHIFENVIGPKMLKXNTRLLVTHSMGYLPOVDYI 839
 DB 780 RAVYSNADIVLFDPLSNVDHAKKHIFENVIGPKMLKXNTRLLVTHSMGYLPOVDYI 839
 QY 840 VMSGKISEMGSYOELLARDGAPAEPLRTYASTEODAEENGTVGSGPKKAKOMENG 899
 DB 840 VMSGKISEMGSYOELLARDGAPAEPLRTYASTEODAEENGTVGSGPKKAKOMENG 899
 QY 900 MLVTDAGKQLOROLSSSSSYSGDISRHNSDALQEAKEKETWKLMEADKQGTQYK 959
 DB 900 MLVTDAGKQLOROLSSSSSYSGDISRHNSDALQEAKEKETWKLMEADKQGTQYK 959
 QY 960 LSVYMDWKKAILGLFISFLSTFLPMCHVYSALASNYMLSLMTDD-PIVNGTEHKVRLSY 1018
 DB 960 LSVYMDWKKAILGLFISFLSTFLPMCHVYSALASNYMLSLMTDDPIVNGTEHKVRLSY 1018
 QY 1019 YGALGISQGLAVFGYSMAVSIQGLASRCLHVDLHSLRSKPSMSPFERTPSGNLVNRSK 1078
 DB 1019 YGALGISQGLAVFGYSMAVSIQGLASRCLHVDLHSLRSKPSMSPFERTPSGNLVNRSK 1078
 QY 1079 ELDTVDKMIPEVYIMFNGSLFNVTGACVILLAPPIAIIIPYGLIYFYQRYVASR 1138
 DB 1079 ELDTVDKMIPEVYIMFNGSLFNVTGACVILLAPPIAIIIPYGLIYFYQRYVASR 1138
 QY 1139 QLRKLESYSRSPVYSHNETLIGSVIRAPEQERFIHQSDLKVDENOKAYPSIVANRW 1198
 DB 1139 QLRKLESYSRSPVYSHNETLIGSVIRAPEQERFIHQSDLKVDENOKAYPSIVANRW 1198
 QY 1199 LAVRLCEVGNCTVFAALFAVISRHSLSAGLVGSYSLOVTTYLNMVMSSEMETNI 1258
 DB 1199 LAVRLCEVGNCTVFAALFAVISRHSLSAGLVGSYSLOVTTYLNMVMSSEMETNI 1258
 QY 1259 VAVRLKESYTEKAPWQIOETAPPSGWPQVGVERRNYTLRREDLDFLRHINVTIN 1318
 DB 1259 VAVRLKESYTEKAPWQIOETAPPSGWPQVGVERRNYTLRREDLDFLRHINVTIN 1318
 QY 1316 GGEKVGIVGRGAKSSLTGLFRINSAGEEIIIDGNIAKIGLHDLRFKITTIPDPV 1375
 DB 1316 GGEKVGIVGRGAKSSLTGLFRINSAGEEIIIDGNIAKIGLHDLRFKITTIPDPV 1375
 QY 1379 LFGSLSLNKNDLPFGQYSDDEEYVTSLELAHLKDQFYSALPDKLHDCAGEGEMLSYGORLY 1438
 DB 1379 LFGSLSLNKNDLPFGQYSDDEEYVTSLELAHLKDQFYSALPDKLHDCAGEGEMLSYGORLY 1438
 QY 1436 CLARALRKTKILVDEATAVDELTDLIQSTIRTOEDCTVLTIAHRLNTINDYTVI 1495
 DB 1436 CLARALRKTKILVDEATAVDELTDLIQSTIRTOEDCTVLTIAHRLNTINDYTVI 1495

RESULT 15
 US-08-462-109A-6
 ; Sequence 6, Application US/08462109A
 ; Patent No. 5882875
 ; GENERAL INFORMATION:
 ; APPLICANT: Cole, Susan P. C.
 ; APPLICANT: Dealey, Roger G.
 ; TITLE OF INVENTION: METHODS FOR IDENTIFYING
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: LAHIVE & COCKFIELD
 ; STREET: 60 State Street, suite 510
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA

ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/462,109A
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/966,923
 FILING DATE: 27-OCT-1992
 APPLICATION NUMBER: 08/029,340
 FILING DATE: 8-MAR-1993
 APPLICATION NUMBER: 08/141,893
 FILING DATE: 26-OCT-1993
 APPLICATION NUMBER: 08/407,207
 FILING DATE: 20-MAR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Decorti, Giulio A. Jr.
 REGISTRATION NUMBER: 31,503
 REFERENCE/DOCKET NUMBER: PQI-002CP4
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1528 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-462-109A-6

Query Match 70.7%; Score 7002.5; DB 1; Length 1528;
 Best Local Similarity 88.0%; Pred. No. 0;
 Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALGFGCAGDSDDLPMNNVMTWNTSNPFTKCFQNTLVVWVPCFLTACGPFYLYLSRH 60
 DB 1 MALSFSCAGDSDDLPMNNVMTWNTSNPFTKCFQNTLVVWVPCFLTACGPFYLYLSRH 60
 QY 61 DRGYIOMTPLNKTALGFLIMIVCMADLFYSFMRSGRIPLAPFVSPFTLIGITLLA 120
 DB 61 DRGIIOMTPLNKTALGFLIMIVCMADLFYSFMRSGRIPLAPFVSPFTLIGITLLA 120
 QY 121 TFLIQLERRKGVQSSGIMLTFMVALVCALAILRSKIMTALKEAQQVDFPDITFYVFS 180
 DB 121 TFLIQLERRKGVQSSGIMLTFMVALVCALAILRSKIMTALKEAQQVDFPDITFYVFS 180
 QY 181 LLLIQVLSCFSDSPFLSETIHDNPNCPSSSASFLSRITTWITGLIVRGYROPLEGSD 240
 DB 181 LVLQVLSCFSDSPFLSETIHDNPNCPSSSASFLSRITTWITGWVHGYROPLEGSD 240
 QY 241 LMSLNKEDTSQQVYVLYKMKKCECAKTRKQPVKVYVS-SKDPDQPKSSSSQDVANEVEA 299
 DB 241 LMSLNKEDTSQQVYVLYKMKKCECAKTRKQPVKVYVS-SKDPDQPKSSSSQDVANEVEA 299
 QY 300 LIVSPKEMNPSEFLKVLKTFGEYFLMSFFFKAIHDLMPFGPQILKLIFVNDYKAP 359
 DB 301 LIVSPKEMNPSEFLKVLKTFGEYFLMSFFFKAIHDLMPFGPQILKLIFVNDYKAP 359
 QY 360 DWQGYFTVLLFVPAACLOTVLHQYFHCYSGMRKIKTAVTGAAYRKALVITNSARISST 419
 DB 361 DWQGYFTVLLFVPAACLOTVLHQYFHCYSGMRKIKTAVTGAAYRKALVITNSARISST 419
 QY 420 VGEIVNLSVDAQGFMDLATYINNIWSAPLOVIALYLLMTNLGPSVLAGAVWVLANVP 479
 DB 421 VGEIVNLSVDAQGFMDLATYINNIWSAPLOVIALYLLMTNLGPSVLAGAVWVLANVP 479
 QY 480 NAYNAKTKTYOVAHMSKDNRIKLMNEILNGIVLKLAYAMELAFKQVLAIRQEBELKVL 539
 DB 481 NAYNAKTKTYOVAHMSKDNRIKLMNEILNGIVLKLAYAMELAFKQVLAIRQEBELKVL 540

QY 540 KKSAYLSAVGFTWVCPBELVALCQFAVYVTTIDENNILLDAQTAFLVALFNILREPLNLL 599
 DB 541 KKSAYLSAVGFTWVCPBELVALCQFAVYVTTIDENNILLDAQTAFLVALFNILREPLNLL 600
 QY 600 PMVSIIVQASVSLKRLRIPLSHELEBDSIERRPVKDGCGTNSITVANAFTWASDPP 659
 DB 601 PMVSIIVQASVSLKRLRIPLSHELEBDSIERRPVKDGCGTNSITVANAFTWASDPP 659
 QY 660 TLNGITSPREGALVAVVGVCCKSSILSALLAEMDKVEGHVALKGSVAVPOQAWION 719
 DB 660 TLNGITSPREGALVAVVGVCCKSSILSALLAEMDKVEGHVALKGSVAVPOQAWION 719
 QY 720 DSLARENILFGQLEPPYRASYIOACALLPDLEIIPSGGRTEIGRGVNLSSGQQRVSILA 779
 DB 720 DSLARENILFGQLEPPYRASYIOACALLPDLEIIPSGGRTEIGRGVNLSSGQQRVSILA 779
 QY 780 RAVYSNADIVLPDPLSAVDAAVGHKHFENYIGPKMLKNTRLIVTHSMGYLPOVDYI 839
 DB 780 RAVYSNADIVLPDPLSAVDAAVGHKHFENYIGPKMLKNTRLIVTHSMGYLPOVDYI 839
 QY 840 VMSGKISSEMSYQELLARDGAPAEPLRTVASTQEOBDEENGVTGVSQPKKAKOMENG 899
 DB 840 VMSGKISSEMSYQELLARDGAPAEPLRTVASTQEOBDEENGVTGVSQPKKAKOMENG 899
 QY 900 MLVYDSAGKOLQROLSSSSSYSQDISRHHNSTALQRAEAKERTWLMENDAKAQGVK 959
 DB 900 MLVYDSAGKOLQROLSSSSSYSQDISRHHNSTALQRAEAKERTWLMENDAKAQGVK 959
 QY 960 LSVYMDWYKAIQGLFISFIFLPMCNHVSALASNYMLSLMTDD-PYNGTOEHKVRSLV 1018
 DB 960 LSVYMDWYKAIQGLFISFIFLPMCNHVSALASNYMLSLMTDD-PYNGTOEHKVRSLV 1018
 QY 956 LSVYMNWYKAIQGLFISFIFLPMCNHVSALASNYMLSLMTDDPPVANGTOANRFRSLV 1015
 DB 956 LSVYMNWYKAIQGLFISFIFLPMCNHVSALASNYMLSLMTDDPPVANGTOANRFRSLV 1015
 QY 1019 YGALGISQGIAPFGYSMAVSIQGIILASRCLHVDLHSLRSKPMSPFERTPSGNLVNRSK 1078
 DB 1019 YGALGISQGIAPFGYSMAVSIQGIILASRCLHVDLHSLRSKPMSPFERTPSGNLVNRSK 1078
 QY 1016 YGALGILQGAALIFGYMAVSIQGIIFASRRLHLDLVNVLRSKPMSPFERTPSGNLVNRSK 1075
 DB 1016 YGALGILQGAALIFGYMAVSIQGIIFASRRLHLDLVNVLRSKPMSPFERTPSGNLVNRSK 1075
 QY 1079 ELDTVDNMIPREVIQMPFGLPNVYGACIVILLATPIAIIIPGLIYFFQRYVASSR 1138
 DB 1079 ELDTVDNMIPREVIQMPFGLPNVYGACIVILLATPIAIIIPGLIYFFQRYVASSR 1138
 QY 1133 QLRLESVSRSPPVSHNETLIGSVIRAFEBQRFTHOSDLKVDENOKAYPSIVANRW 1198
 DB 1133 QLRLESVSRSPPVSHNETLIGSVIRAFEBQRFTHOSDLKVDENOKAYPSIVANRW 1198
 QY 1136 QLRLESVSRSPPVSHNETLIGSVIRAFEBQRFTHOSDLKVDENOKAYPSIVANRW 1195
 DB 1136 QLRLESVSRSPPVSHNETLIGSVIRAFEBQRFTHOSDLKVDENOKAYPSIVANRW 1195
 QY 1199 LAVRLCEVGNCTVLFALFAVISRHSLSAGLVGSVSLQVTTYLNMLVRMSSEMETNI 1258
 DB 1199 LAVRLCEVGNCTVLFALFAVISRHSLSAGLVGSVSLQVTTYLNMLVRMSSEMETNI 1258
 QY 1259 VAVRRLKEYSBTEKEAWOIOETNAPBSMPOGVHVEFRNTCLARYEDLDVLRHINVTIN 1318
 DB 1259 VAVRRLKEYSBTEKEAWOIOETNAPBSMPOGVHVEFRNTCLARYEDLDVLRHINVTIE 1315
 QY 1319 GGEKVGIVGRTGAGKSLTGLPRINSAGEEIIIDGINIAKIGLHDLRFKTIIPDPV 1378
 DB 1319 GGEKVGIVGRTGAGKSLTGLPRINSAGEEIIIDGINIAKIGLHDLRFKTIIPDPV 1378
 QY 1316 GGEKVGIVGRTGAGKSLTGLPRINSAGEEIIIDGINIAKIGLHDLRFKTIIPDPV 1375
 DB 1316 GGEKVGIVGRTGAGKSLTGLPRINSAGEEIIIDGINIAKIGLHDLRFKTIIPDPV 1375
 QY 1379 LFGSGLRMLNDPFSQYSDBEWNTSLBLAHLKDFVSAJLBDLDEHCAEGENLSVGQROLV 1438
 DB 1379 LFGSGLRMLNDPFSQYSDBEWNTSLBLAHLKDFVSAJLBDLDEHCAEGENLSVGQROLV 1435
 QY 1439 CLARALIRKTKIIVLDEATAVVDLETDLIQSTIRIQFEDCTVLTIAHRLNTIMDYTRVI 1498
 DB 1439 CLARALIRKTKIIVLDEATAVVDLETDLIQSTIRIQFEDCTVLTIAHRLNTIMDYTRVI 1495
 QY 1499 VLDKGEIOEYGAPEBDLQQRGLFFYSMAKDAQGLV 1531
 DB 1499 VLDKGEIOEYGAPEBDLQQRGLFFYSMAKDAQGLV 1528

Search completed: December 15, 2005, 15:49:51
 Job time : 48.0811 secs

Blank (uspb)

QY 361 WQGFYTLVLTAVTACLOTLVHGHYHICFVSGMRKIKTAVIGAVRKALVTINSARKSSTV 420
DB 361 WQGFYTLVLTAVTACLOTLVHGHYHICFVSGMRKIKTAVIGAVRKALVTINSARKSSTV 420
QY 421 GEIVNLSVDAHQREMDLATYINMISAPLOVTLAYLMLNIGBSVLAGVAVMLVMVYN 480
DB 421 GEIVNLSVDAHQREMDLATYINMISAPLOVTLAYLMLNIGBSVLAGVAVMLVMVYN 480
QY 481 AVMAKTKTYOVAMHKSNDRIKLMNEILNGIKVLKLYAMELAFKDYLAIRQELKYLK 540
DB 481 AVMAKTKTYOVAMHKSNDRIKLMNEILNGIKVLKLYAMELAFKDYLAIRQELKYLK 540
QY 541 KSAVISAAGVTWCTPPLVALCTPAVYVTTDENNILDQAFAVSLAFNLIRPLNITLP 600
DB 541 KSAVISAAGVTWCTPPLVALCTPAVYVTTDENNILDQAFAVSLAFNLIRPLNITLP 600
QY 601 MVISIVASVSLKRLIFLSHEELEPDSIERRPVKGDTGNTSITVRNATFTWASDPT 660
DB 601 MVISIVASVSLKRLIFLSHEELEPDSIERRPVKGDTGNTSITVRNATFTWASDPT 660
QY 661 LNGITFSIPREGALVAVGVQVCGKSSLSALLAEMDKYEGHVAIKGSVAYVPOQAMIOND 720
DB 661 LNGITFSIPREGALVAVGVQVCGKSSLSALLAEMDKYEGHVAIKGSVAYVPOQAMIOND 720
QY 721 SLRENIIFGCOLLEBPYRSVIOACALPDLEITLPSGDRTEIGEKVNLSGQOKORVSLAR 780
DB 721 SLRENIIFGCOLLEBPYRSVIOACALPDLEITLPSGDRTEIGEKVNLSGQOKORVSLAR 780
QY 781 AVYNADITLFDLPLSADAHVGHKIPENVIGPKMKLNKTRILVTHSMNLTPOVDVITV 840
DB 781 AVYNADITLFDLPLSADAHVGHKIPENVIGPKMKLNKTRILVTHSMNLTPOVDVITV 840
QY 841 MSGKISMGYSVOELLARDGAFAEFLRTYASTEOQDAENGCVTVSGPGKEAKOMENGM 900
DB 841 MSGKISMGYSVOELLARDGAFAEFLRTYASTEOQDAENGCVTVSGPGKEAKOMENGM 900
QY 901 LVTDGAKOLQROUSSSSSYSGDISRHHNSTAELOKAEKKEETWKLMEADKAQOTQVYL 960
DB 901 LVTDGAKOLQROUSSSSSYSGDISRHHNSTAELOKAEKKEETWKLMEADKAQOTQVYL 960
QY 961 SVYNDYMAIGLFTSFLSIFLMCHNVASALSNYMLSLMTDDPIYNGOEHKTVLSTYVG 1020
DB 961 SVYNDYMAIGLFTSFLSIFLMCHNVASALSNYMLSLMTDDPIYNGOEHKTVLSTYVG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGGILASRCHVDLHSLRSPMSFFERTSPSGMLVNRFSKEL 1080
DB 1021 ALGISOGIAVFGYSMAVSIIGGILASRCHVDLHSLRSPMSFFERTSPSGMLVNRFSKEL 1080
QY 1081 DTVDSMIPEVIKMFWSLFFNVIAGACIVILLATPIAAIIPPLGLIYFFVQRPYVASSRQL 1140
DB 1081 DTVDSMIPEVIKMFWSLFFNVIAGACIVILLATPIAAIIPPLGLIYFFVQRPYVASSRQL 1140
QY 1141 KRLESVSRSVYVSHFNEITLGVSVIRAFEEQERFHODLKVDEOKAYYSIVANRWLA 1200
DB 1141 KRLESVSRSVYVSHFNEITLGVSVIRAFEEQERFHODLKVDEOKAYYSIVANRWLA 1200
QY 1201 VRLECVGNCTVLPALFAVISRHSLSAGLVGSLVSYSIQVTTYLWMLVRMSSEMETNIVA 1260
DB 1201 VRLECVGNCTVLPALFAVISRHSLSAGLVGSLVSYSIQVTTYLWMLVRMSSEMETNIVA 1260
QY 1261 VERLKEVSETEKEAPWQIOETAPSPSWPOGRVEFRNCTARYREDDLVLAHINVTINGG 1320
DB 1261 VERLKEVSETEKEAPWQIOETAPSPSWPOGRVEFRNCTARYREDDLVLAHINVTINGG 1320
QY 1321 EKVGIVERTGAKGSSLTGLFRINESAGEHIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
DB 1321 EKVGIVERTGAKGSSLTGLFRINESAGEHIIIDGINIAKIGLHDLRFKTIITIPDPVLF 1380
QY 1381 SSGSRMNLDPFSPQSDSEVWTSLELAHLKDFVSAIPDLKDHBCAGGNNLSVGGQQLVCL 1440
DB 1381 SSGSRMNLDPFSPQSDSEVWTSLELAHLKDFVSAIPDLKDHBCAGGNNLSVGGQQLVCL 1440

QY 1441 APALIRKTKIIVDBANAVDLFTDDLIQSTIRFOFEDCTVLTIAHRLNTIMDYTRIVL 1500
DB 1441 APALIRKTKIIVDBANAVDLFTDDLIQSTIRFOFEDCTVLTIAHRLNTIMDYTRIVL 1500
QY 1501 DKGEIOEYGAUSDLLQORGLFYSMAKDAVLGGGGGGLSKRGIIPEEYVLTRLAEDPAE 1560
DB 1501 DKGEIOEYGAUSDLLQORGLFYSMAKDAVLGGGGGGLSKRGIIPEEYVLTRLAEDPAE 1560
QY 1561 PRYTRERRARFVSKGNCNVAHKNIREQGRFLQDVFTTLVDLKMPTLLIFTMSFLCSW 1620
DB 1561 PRYTRERRARFVSKGNCNVAHKNIREQGRFLQDVFTTLVDLKMPTLLIFTMSFLCSW 1620
QY 1621 LLPAMWMLIAFAHGDIAPEGSTNVCVTSIHSFSSAFLFSIEVQVTIGFGRMVTEBCP 1680
DB 1621 LLPAMWMLIAFAHGDIAPEGSTNVCVTSIHSFSSAFLFSIEVQVTIGFGRMVTEBCP 1680
QY 1681 LAILIIVQNTVGMINAMIGCI FMKTAQHRAEFTLIFSKHAVITLRHGLCPMLRPG 1740
DB 1681 LAILIIVQNTVGMINAMIGCI FMKTAQHRAEFTLIFSKHAVITLRHGLCPMLRPG 1740
QY 1741 DLKRSMTISATIHQVVRKTTSPGEVVPPLHQVDIPMENGVGNGIFLVAELIYHVIDS 1800
DB 1741 DLKRSMTISATIHQVVRKTTSPGEVVPPLHQVDIPMENGVGNGIFLVAELIYHVIDS 1800
QY 1801 NSPLYDLAPSDLHHODLEIIVILEGVETTGITTOARTSYLADEIIMGQRFVPIVAED 1860
DB 1801 NSPLYDLAPSDLHHODLEIIVILEGVETTGITTOARTSYLADEIIMGQRFVPIVAED 1860
QY 1861 GRYSVDYSKFGNTLKVPLPCTARQLEDBRSLDALTLASRGLBAASVAVAKKPFES 1920
DB 1861 GRYSVDYSKFGNTLKVPLPCTARQLEDBRSLDALTLASRGLBAASVAVAKKPFES 1920
QY 1921 ISPDLS 1927
DB 1921 ISPDLS 1927

RESULT 2
US-10-665-283-1
; Sequence 1, Application US/10665283
; Publication No. US20050063989A1
; GENERAL INFORMATION:
; APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
; APPLICANT: DERAND, Renaud
; APPLICANT: GARCIA, Elisabeth
; APPLICANT: PROST, Anne-Lise
; APPLICANT: REVILLAUD, Jean
; APPLICANT: VIVAUDOU, Michel
; TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF
; FILE REFERENCE: P263US104
; CURRENT APPLICATION NUMBER: US/10/665,283
; CURRENT FILING DATE: 2003-09-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1927
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-665-283-1

Query Match 99.9%; Score 9891; DB 5; Length 1927;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1925; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFSGADGSDPLMDMNTWTNTSNDFTKCFONTYLVWVPCYYLWACFPFFYLYLSRH 60
DB 1 MALRGFSGADGSDPLMDMNTWTNTSNDFTKCFONTYLVWVPCYYLWACFPFFYLYLSRH 60
QY 61 DRGIQMTPLNKTKTALGFLIMTICWADLFPSFERSRGFLAVVPFVSPFLIGITTLA 120
DB 61 DRGIQMTPLNKTKTALGFLIMTICWADLFPSFERSRGFLAVVPFVSPFLIGITTLA 120
QY 121 TFLQLERRKGVQSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDLFRDITFYVYS 180

```

Db      121  TFLQLRRKRGVSSGIMLTFMLVALVCAALILRSKIMTALKEDAOVDFRDIIFYYFS 180
Qy      181  LLLQLVLSCSDSPLESETIHDNCPRESSASFLSITFWMTGLIVRGYRPLSGD 240
Db      181  LLLQLVLSCSDSPLESETIHDNCPRESSASFLSITFWMTGLIVRGYRPLSGD 240
Qy      241  LMSLNKEDTSBOVPVLVKNWKECAKTRKQPVKVYSSKDPAPCKESSKVDANEVEAL 300
Db      241  LMSLNKEDTSBOVPVLVKNWKECAKTRKQPVKVYSSKDPAPCKESSKVDANEVEAL 300
Qy      301  IVSPQKEMNPSLFKVLYKTFGPFLMSFFPKAHDLMFSGPOLIKLIFVNDTRYAPD 360
Db      301  IVSPQKEMNPSLFKVLYKTFGPFLMSFFPKAHDLMFSGPOLIKLIFVNDTRYAPD 360
Qy      361  MOGFYFVLLFVTAQLOTLVLAHQFPHICFVSGMRTKAVIGAVYRKALVTNSARKSSTV 420
Db      361  MOGFYFVLLFVTAQLOTLVLAHQFPHICFVSGMRTKAVIGAVYRKALVTNSARKSSTV 420
Qy      421  GEIYNLMSVDAQRFMDLATYINMISAPLOVITLALYLMLNIGPSVLAGVAVWMLAMPV 480
Db      421  GEIYNLMSVDAQRFMDLATYINMISAPLOVITLALYLMLNIGPSVLAGVAVWMLAMPV 480
Qy      481  AVAMAKTKTYQVAHMKSKDNRIKLMNEILINGIKVLKYLAMELAFKDKVLAIROBELVYK 540
Db      481  AVAMAKTKTYQVAHMKSKDNRIKLMNEILINGIKVLKYLAMELAFKDKVLAIROBELVYK 540
Qy      541  KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNIIDAQTAFSLALFNLIRPPLILP 600
Db      541  KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNIIDAQTAFSLALFNLIRPPLILP 600
Qy      601  MVYSIIQASVSLKRLRIFLSHEELBPSIERRPVKOGGNGSITVANAFTWARSBPT 660
Db      601  MVYSIIQASVSLKRLRIFLSHEELBPSIERRPVKOGGNGSITVANAFTWARSBPT 660
Qy      661  LMGITFESIPGALVAVVGVGCCGKSLLSALLAEMDYEGVNAIKGSVAVYPOQAWIOND 720
Db      661  LMGITFESIPGALVAVVGVGCCGKSLLSALLAEMDYEGVNAIKGSVAVYPOQAWIOND 720
Qy      721  SLRENIILFGCOLBEPYRYSYIOACALLPDLBILPDSGRTEIGEKNVLSGGQOKRVSLAR 780
Db      721  SLRENIILFGCOLBEPYRYSYIOACALLPDLBILPDSGRTEIGEKNVLSGGQOKRVSLAR 780
Qy      781  AAVSNADITYLPDPLSAVDHVGKHFENYIGPKMKMKNTRILLVHSMGYLPOVDYIIV 840
Db      781  AAVSNADITYLPDPLSAVDHVGKHFENYIGPKMKMKNTRILLVHSMGYLPOVDYIIV 840
Qy      841  MSGGKISEMGSYOELLARDGAFAEFLRTYASTEQEODAEENGVTGVSQPKGAOKOMENG 900
Db      841  MSGGKISEMGSYOELLARDGAFAEFLRTYASTEQEODAEENGVTGVSQPKGAOKOMENG 900
Qy      901  LVYDSAGKQLOROLSSSSSYSGDISRHNSSTAELIOKAKEETWKMMEADKAQTQVYL 960
Db      901  LVYDSAGKQLOROLSSSSSYSGDISRHNSSTAELIOKAKEETWKMMEADKAQTQVYL 960
Qy      961  SVYMDTWKALGTLFSLSTFLFMCNHSALASNYWLSLWDDPILVNGTQHTKVRSLVYG 1020
Db      961  SVYMDTWKALGTLFSLSTFLFMCNHSALASNYWLSLWDDPILVNGTQHTKVRSLVYG 1020
Qy      1021  ALGISOGIAVFGYSMAVISIGIILASRCLAYDLHLSILRSFMSFEERTPSGNLVNRFSEKL 1080
Db      1021  ALGISOGIAVFGYSMAVISIGIILASRCLAYDLHLSILRSFMSFEERTPSGNLVNRFSEKL 1080
Qy      1081  DTYDSMIPYIKMFMSGLFNVIIGACIVILLATPAAIIIPPLGIYFFVGRFYASSROL 1140
Db      1081  DTYDSMIPYIKMFMSGLFNVIIGACIVILLATPAAIIIPPLGIYFFVGRFYASSROL 1140
Qy      1141  KRLSVSRSPVYSHFNMTLLGVSVIRAFEEOEERFIHOSDLKVDENQAAVYPSIVANKWLA 1200
Db      1141  KRLSVSRSPVYSHFNMTLLGVSVIRAFEEOEERFIHOSDLKVDENQAAVYPSIVANKWLA 1200
Qy      1201  VRLCVCNCIVLFAALFAVISRHSLSAGLVGLSVYSYLOVTTYLNMVLVRMSSEMETNIVA 1260

```

```

Db      1201  VRLCVCNCIVLFAALFAVISRHSLSAGLVGLSVYSYLOVTTYLNMVLVRMSSEMETNIVA 1260
Qy      1261  VERLKEYSETEKEAPMOIOSTAPSSWPQVGRVFNRYCLRYRDDLDPVLRHINVTNGG 1320
Db      1261  VERLKEYSETEKEAPMOIOSTAPSSWPQVGRVFNRYCLRYRDDLDPVLRHINVTNGG 1320
Qy      1321  EKVGIVGRTAGKSSSLTGLFPRINESAGEIILIDGINIAKIGLDLRFKTIIFIQDDPLF 1380
Db      1321  EKVGIVGRTAGKSSSLTGLFPRINESAGEIILIDGINIAKIGLDLRFKTIIFIQDDPLF 1380
Qy      1381  SGLRMLNLPFSQYSDSEWVTSLELAHKDPVSALPDKLDHECAGEGENTSVGQROLVCL 1440
Db      1381  SGLRMLNLPFSQYSDSEWVTSLELAHKDPVSALPDKLDHECAGEGENTSVGQROLVCL 1440
Qy      1441  ARALRRTKTLVNDDEATAVADLETDLDLOSTIRQFEDCTVLTAAHLNTIMOTRYIVL 1500
Db      1441  ARALRRTKTLVNDDEATAVADLETDLDLOSTIRQFEDCTVLTAAHLNTIMOTRYIVL 1500
Qy      1501  DKGEIOEYGAPSDLOOQGLFYSMAXDAGLVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
Db      1501  DKGEIOEYGAPSDLOOQGLFYSMAXDAGLVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
Qy      1561  PRYTRERRARFVSKKGCNCVNAHKNIREQRFLQDVFTTLVDLKMPTLLIFTWSFLCSW 1620
Db      1561  PRYTRERRARFVSKKGCNCVNAHKNIREQRFLQDVFTTLVDLKMPTLLIFTWSFLCSW 1620
Qy      1621  LFLPAMWMLIAFAHGLDAPBEGTNVPCVTSIHSSSAFLSIEVOVTTIGRGRMVTBEP 1680
Db      1621  LFLPAMWMLIAFAHGLDAPBEGTNVPCVTSIHSSSAFLSIEVOVTTIGRGRMVTBEP 1680
Qy      1681  LALILILIVONIYGIMINAMLGCI FMKTAQAHRAETLISKHAIVTLRHRGRLCEMRLVG 1740
Db      1681  LALILILIVONIYGIMINAMLGCI FMKTAQAHRAETLISKHAIVTLRHRGRLCEMRLVG 1740
Qy      1741  DLKRSMLISATIMHQRVKTSPBEGEVPLHQVDIPMENGVGNGIFLVAPLIIHYIDS 1800
Db      1741  DLKRSMLISATIMHQRVKTSPBEGEVPLHQVDIPMENGVGNGIFLVAPLIIHYIDS 1800
Qy      1801  NSPLYDLAPSDLHHODLEIIVLEGEVETTGITTOARTSYLADEIIMGQRFVPIVAED 1860
Db      1801  NSPLYDLAPSDLHHODLEIIVLEGEVETTGITTOARTSYLADEIIMGQRFVPIVAED 1860
Qy      1861  GRYSVDYSKRGNTIKVPTPLCTAROLDDEDSLDDALLTLASRGLPAAVAVAKRKS 1920
Db      1861  GRYSVDYSKRGNTIKVPTPLCTAROLDDEDSLDDALLTLASRGLPAAVAVAKRKS 1920
Qy      1921  ISPDLS 1927
Db      1921  ISPDLS 1927

RESULT 3
US-10-665-283-8
; Sequence 8, Application US/10665283
; Publication No. US20050063989A1
; GENERAL INFORMATION:
; APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
; APPLICANT: DERAND, Renaud
; APPLICANT: GARCIA, Elisabeth
; APPLICANT: PROST, Anne-Lise
; APPLICANT: REVILLION, Jean
; APPLICANT: VIVAUDOU, Michel
; TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF
; FILE REFERENCE: F263US104
; CURRENT APPLICATION NUMBER: US/10/665,283
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 1947
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-665-283-8

```

Query Match 99.6%; Score 9861; DB 5; Length 1947;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 1925; Conservative 0; Mismatches 2; Indels 20; Gaps 2;

```

Qy 1 MALRFGSADSDLDLMDMNTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFFLYLSRH 60
Db 1 MALRFGSADSDLDLMDMNTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFFLYLSRH 60
Qy 61 DRGIQMTPLNKKTALGELLMTICWADLFYSFMRSRGIFLAPVFLVSPLLGLITLLA 120
Db 61 DRGIQMTPLNKKTALGELLMTICWADLFYSFMRSRGIFLAPVFLVSPLLGLITLLA 120
Qy 121 TFLIOLERRKGVSSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDLDFRDIFFYVFS 180
Db 121 TFLIOLERRKGVSSGIMLTFMLVALVCALAILRSKIMTALKEDAQVDLDFRDIFFYVFS 180
Qy 181 LLLIQLVLSCFSDNSPLFSETIHDNPPCESSASFLSKITFWMTGLIVRGYROPLEGSD 240
Db 181 LLLIQLVLSCFSDNSPLFSETIHDNPPCESSASFLSKITFWMTGLIVRGYROPLEGSD 240
Qy 241 LMSLNKEDTSRQVVPVLVKNMKKECAKTRKQPVKVYVSSKDPAPQKSSKVDANEVEAL 300
Db 241 LMSLNKEDTSRQVVPVLVKNMKKECAKTRKQPVKVYVSSKDPAPQKSSKVDANEVEAL 300
Qy 301 IVKSPOKEMNBSLFKVLYKTFGPYFLMSFFKAIHDLMMFSGPOLIKLLIFVNDTKAPD 360
Db 301 IVKSPOKEMNBSLFKVLYKTFGPYFLMSFFKAIHDLMMFSGPOLIKLLIFVNDTKAPD 360
Qy 361 WQGFYTVLLFVYACLQTLVLAHQYFHICFVSGMRIKTAVIGAVYRKALVITNSARKSTV 420
Db 361 WQGFYTVLLFVYACLQTLVLAHQYFHICFVSGMRIKTAVIGAVYRKALVITNSARKSTV 420
Qy 421 GEIYNLMSVDORFMDLTYINMISAPLOYLIALYLLMLNGSPVLGAVMVLVNVN 480
Db 421 GEIYNLMSVDORFMDLTYINMISAPLOYLIALYLLMLNGSPVLGAVMVLVNVN 480
Qy 481 AVAMAMTKTYOVANHKSKDNRIKLMNEILNGIKVLKLYAMELAFKDYALROBELKYLK 540
Db 481 AVAMAMTKTYOVANHKSKDNRIKLMNEILNGIKVLKLYAMELAFKDYALROBELKYLK 540
Qy 541 KSAVLSAVGTWCTPPLVALCTPAVYVTTDENNILDQAOTAFVSLAFNLLRPLNLTLP 600
Db 541 KSAVLSAVGTWCTPPLVALCTPAVYVTTDENNILDQAOTAFVSLAFNLLRPLNLTLP 600
Qy 601 MVISIVASVSLKRLRIFLSHEELEPDSIBRRPVKDGSTNSITVRNATFTWASDPT 660
Db 601 MVISIVASVSLKRLRIFLSHEELEPDSIBRRPVKDGSTNSITVRNATFTWASDPT 660
Qy 661 LNGITFSPREGALVAVVQVCGSKSLSLALAEADKYGVAIKGSVAVYVPOQAMIOND 720
Db 661 LNGITFSPREGALVAVVQVCGSKSLSLALAEADKYGVAIKGSVAVYVPOQAMIOND 720
Qy 721 SLRENIILFCQLEEBYRSVIQACALLPDLLETLPDSGRTEIGEKVNLSGQOKORVSLAR 780
Db 721 SLRENIILFCQLEEBYRSVIQACALLPDLLETLPDSGRTEIGEKVNLSGQOKORVSLAR 780
Qy 781 AVYSNADITYLPDDPLSAVDAAHVKHIFENVIQPKMLNKRILVTHSMSTYLPQVDVITV 840
Db 781 AVYSNADITYLPDDPLSAVDAAHVKHIFENVIQPKMLNKRILVTHSMSTYLPQVDVITV 840
Qy 841 MSGGKISEMGSYOELLARDGAFAELRTYASTEQDADAEENGVTGVSQKGEAKOMENGM 900
Db 841 MSGGKISEMGSYOELLARDGAFAELRTYASTEQDADAEENGVTGVSQKGEAKOMENGM 900
Qy 901 LVYDSAGQOLQOLSSSSSYSGDISRHNSTALQKAEAKKETWKLMEADRAQOTGOYL 960
Db 901 LVYDSAGQOLQOLSSSSSYSGDISRHNSTALQKAEAKKETWKLMEADRAQOTGOYL 960
Qy 961 SVYWDYMAKIGLIFSLIFLPMCHVSAALASNYLSTMTDPIYNGOEHKRYLSTYSG 1020
Db 961 SVYWDYMAKIGLIFSLIFLPMCHVSAALASNYLSTMTDPIYNGOEHKRYLSTYSG 1020

```

```

Qy 1021 ALGISQIAVFGYSMAVSGIILASRCLHYDLHSILRSFMSFPERTPSGNLVNRBEKE 1080
Db 1021 ALGISQIAVFGYSMAVSGIILASRCLHYDLHSILRSFMSFPERTPSGNLVNRBEKE 1080
Qy 1081 DTVDSMTPEVITKPMGSLFENVIGACTVITLATAPIAIIIPPLGIYFPVQRFVYASSROL 1140
Db 1081 DTVDSMTPEVITKPMGSLFENVIGACTVITLATAPIAIIIPPLGIYFPVQRFVYASSROL 1140
Qy 1141 KRLSVSRSPVYSHFNTLLGVSYIRAFEEQERFIHOSDLKVDENQAKYPSIVANWMLA 1200
Db 1141 KRLSVSRSPVYSHFNTLLGVSYIRAFEEQERFIHOSDLKVDENQAKYPSIVANWMLA 1200
Qy 1201 VRLCEVNCIYLPALPAVIRSRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSSEMETNIVA 1260
Db 1201 VRLCEVNCIYLPALPAVIRSRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSSEMETNIVA 1260
Qy 1261 VERLKEYSETEKEPMQIQTAPSSWPQVGRVFRNYCLRYREDLDFVLPHINVTINGG 1320
Db 1261 VERLKEYSETEKEPMQIQTAPSSWPQVGRVFRNYCLRYREDLDFVLPHINVTINGG 1320
Qy 1321 EKVGIYGRGTAGKSLTLGLFRINESAGEIITIDGINIAKIGLHDLRFKTIIPQDPVL 1380
Db 1321 EKVGIYGRGTAGKSLTLGLFRINESAGEIITIDGINIAKIGLHDLRFKTIIPQDPVL 1380
Qy 1381 SGLRNMULDPSQYSDDEWVTSLELAHLKDFVSALPKLDHECAEGENLSVGROQLVCL 1440
Db 1381 SGLRNMULDPSQYSDDEWVTSLELAHLKDFVSALPKLDHECAEGENLSVGROQLVCL 1440
Qy 1441 ARALLRTKTLIVLDEAFAVNDLETFDDLIQSTIRPOFEDCVTLTAHRLNTMDTRYVL 1500
Db 1441 ARALLRTKTLIVLDEAFAVNDLETFDDLIQSTIRPOFEDCVTLTAHRLNTMDTRYVL 1500
Qy 1501 DKGBIQRXYGAPSDLLQORGLFYYSMAKDAGLVGGGGGMLSRKGIIPPEYVLTRLAEDPAE 1560
Db 1501 DKGBIQRXYGAPSDLLQORGLFYYSMAKDAGLVGGGGGMLSRKGIIPPEYVLTRLAEDPAE 1560
Qy 1561 PRYTRERRARFVSKKNCNVAAHKNIREQGRFLQDVFVTLVDLKPHTLLIFTWSFLCSW 1620
Db 1561 PRYTRERRARFVSKKNCNVAAHKNIREQGRFLQDVFVTLVDLKPHTLLIFTWSFLCSW 1620
Qy 1621 LLFAMWMLLAFAG-----DLAP-----GEGTNVPCVTSIHSSSAFLF 1660
Db 1621 LLFAMWMLLAFAG-----DLAP-----GEGTNVPCVTSIHSSSAFLF 1660
Qy 1661 SIEVQVITGFGGRWTECEPLAILILIVONIVGLMINAIMGCIEMKTAQAHRAETLIF 1720
Db 1661 SIEVQVITGFGGRWTECEPLAILILIVONIVGLMINAIMGCIEMKTAQAHRAETLIF 1720
Qy 1721 SKHAVITLRHGRCLFMLRVGDLRKSMTISATIHQVVRKTTSPBGEVVPPLHQVDIPMENG 1780
Db 1721 SKHAVITLRHGRCLFMLRVGDLRKSMTISATIHQVVRKTTSPBGEVVPPLHQVDIPMENG 1780
Qy 1781 VGGNGIFLVAELIYYHYIDSNSPLDYDLAPSLHHRHOLEIIVILEGVETTGITTOARTS 1840
Db 1781 VGGNGIFLVAELIYYHYIDSNSPLDYDLAPSLHHRHOLEIIVILEGVETTGITTOARTS 1840
Qy 1801 VGGNGIFLVAELIYYHYIDSNSPLDYDLAPSLHHRHOLEIIVILEGVETTGITTOARTS 1860
Db 1801 VGGNGIFLVAELIYYHYIDSNSPLDYDLAPSLHHRHOLEIIVILEGVETTGITTOARTS 1860
Qy 1841 YLADEIILMGQRFVIVABEDGRYSVDYSKFCNTIKVPTPLCTARQOLDEDRSLDALTLAS 1900
Db 1841 YLADEIILMGQRFVIVABEDGRYSVDYSKFCNTIKVPTPLCTARQOLDEDRSLDALTLAS 1900
Qy 1901 SRGPLRAASVAVAKAKPFSISPDSLS 1927
Db 1901 SRGPLRAASVAVAKAKPFSISPDSLS 1927

```

RESULT 4
 US-10-665-283-4
 ; Sequence 4, Application US/10665283
 ; Publication No. US20050063989A1
 ; GENERAL INFORMATION:
 ; APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
 ; APPLICANT: DERAND, Renaud
 ; APPLICANT: GARCIA, Elisabeth

```

; APPLICANT: PROST, Anne-lise
; APPLICANT: REVILLOUP, Jean
; APPLICANT: VIVAUOU, Michel
; TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF
; FILE REFERENCE: P63US104
; CURRENT APPLICATION NUMBER: US/10/665, 283
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1891
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-665-283-4

Query Match      98.3%; Score 9734; DB 5; Length 1891;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MALRGFCADGSDPLMDMNTMTNTPDKCFQNTLVWVPCFYLMACPFYFLYLSRH 60
DB      1 MALRGFCADGSDPLMDMNTMTNTPDKCFQNTLVWVPCFYLMACPFYFLYLSRH 60

QY      61 DRGYIQMTPLNKTALGFLMIVCMADLFYSFWEBSRGIFLAVFLVSPTLGITTLLA 120
DB      61 DRGYIQMTPLNKTALGFLMIVCMADLFYSFWEBSRGIFLAVFLVSPTLGITTLLA 120

QY      121 TFLIOLERRKGVOSGIMLTFWLVALCALALRSKMTALKEDAQVDFRDIIFYVFS 180
DB      121 TFLIOLERRKGVOSGIMLTFWLVALCALALRSKMTALKEDAQVDFRDIIFYVFS 180

QY      181 LLLIOLVLSGCFSDRSPFSETHDNPCESSASFISRTFPMWTTGLIVGYQPLEGSD 240
DB      181 LLLIOLVLSGCFSDRSPFSETHDNPCESSASFISRTFPMWTTGLIVGYQPLEGSD 240

QY      241 LMSLNKEDTSEQVVPVLYKMKKECAKTRKQPVVYVSSKDPQPKSSKVDANEVEAL 300
DB      241 LMSLNKEDTSEQVVPVLYKMKKECAKTRKQPVVYVSSKDPQPKSSKVDANEVEAL 300

QY      301 IVSPKQKMPNLSFKVLYKTFGPFYLMSPFKAIHDLMFSGPOLIKLIIKPVNDIYAPD 360
DB      301 IVSPKQKMPNLSFKVLYKTFGPFYLMSPFKAIHDLMFSGPOLIKLIIKPVNDIYAPD 360

QY      361 WQGFYVVLVFAVACIOTLVLHOVFI CFVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420
DB      361 WQGFYVVLVFAVACIOTLVLHOVFI CFVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420

QY      421 GEIVNLMSVDAORFMDLATYINMIWSAPLQVILALYLLMLNTLGPVSLAGVAVWLVMPVN 480
DB      421 GEIVNLMSVDAORFMDLATYINMIWSAPLQVILALYLLMLNTLGPVSLAGVAVWLVMPVN 480

QY      481 AVNAMKTKTYQVAHMSKDNRIKLMNEILINGIKVLYKLYAMELAFKDVLAIRBELKVLK 540
DB      481 AVNAMKTKTYQVAHMSKDNRIKLMNEILINGIKVLYKLYAMELAFKDVLAIRBELKVLK 540

QY      541 KSAVLSAVGFTVWCPTPELVALCTPAVYVITIDENNIIDAOPTAVSLAEHILAFPLNLP 600
DB      541 KSAVLSAVGFTVWCPTPELVALCTPAVYVITIDENNIIDAOPTAVSLAEHILAFPLNLP 600

QY      601 MVISSIVQASVSLKRLIFLSHEELBPSIERRPVKQGGTNSITVNAATFTVARSDEPT 660
DB      601 MVISSIVQASVSLKRLIFLSHEELBPSIERRPVKQGGTNSITVNAATFTVARSDEPT 660

QY      661 LMGITTSIPGALVAVVGQVCGCKSSLSALLAEMDVKEGHVAIKGSVAVVPQAMVQND 720
DB      661 LMGITTSIPGALVAVVGQVCGCKSSLSALLAEMDVKEGHVAIKGSVAVVPQAMVQND 720

QY      721 STRENTILFGQLEPEPYRSVIOACALPDLEIILPSGGRTEIGKGVNLGGGQORVSLAR 780
DB      721 STRENTILFGQLEPEPYRSVIOACALPDLEIILPSGGRTEIGKGVNLGGGQORVSLAR 780

QY      781 AVYSNADIVLFDPLSLAVDAVGHKIIFENVIGPKMKLNKTRLLVTHSMSYLPQVDVIY 840
DB      781 AVYSNADIVLFDPLSLAVDAVGHKIIFENVIGPKMKLNKTRLLVTHSMSYLPQVDVIY 840
```

```

DB      781 AVYSNADIVLFDPLSLAVDAVGHKIIFENVIGPKMKLNKTRLLVTHSMSYLPQVDVIY 840
QY      841 MSGGKISEMSXOELLARDQAPAEPLRTVASTBOEOAEENGTVGSGPKKAKOMNGM 900
DB      841 MSGGKISEMSXOELLARDQAPAEPLRTVASTBOEOAEENGTVGSGPKKAKOMNGM 900

QY      901 LVYDSAGKQOROLSSSSSYSGDISRHHNSTAELOKAEKKEFTWKLMEADKQOTGVKL 960
DB      901 LVYDSAGKQOROLSSSSSYSGDISRHHNSTAELOKAEKKEFTWKLMEADKQOTGVKL 960

QY      961 SVYWDYWKALGLFISFLISFLFMCNHYSAALASNYMSLWTDPIVNGTOEHTVRLSVYG 1020
DB      961 SVYWDYWKALGLFISFLISFLFMCNHYSAALASNYMSLWTDPIVNGTOEHTVRLSVYG 1020

QY      1021 ALGISQIAVFGYSMAVSTGIIASRCLHVDLHSLIRSMSPFEERPSGNLVNREKEL 1080
DB      1021 ALGISQIAVFGYSMAVSTGIIASRCLHVDLHSLIRSMSPFEERPSGNLVNREKEL 1080

QY      1081 DTVDSMIPEVIKMFMSGLFNVIGACIYLLATPAAIIIPPLGIYFFVORFYVASSROL 1140
DB      1081 DTVDSMIPEVIKMFMSGLFNVIGACIYLLATPAAIIIPPLGIYFFVORFYVASSROL 1140

QY      1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDUKVDENOKAYVPSIVANRWLA 1200
DB      1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDUKVDENOKAYVPSIVANRWLA 1200

QY      1201 VRLBCVNCIVLFAALPAVTSRHSLSAGLVGSVSVSLQVTTTILNMLVMSSEMETNIVA 1260
DB      1201 VRLBCVNCIVLFAALPAVTSRHSLSAGLVGSVSVSLQVTTTILNMLVMSSEMETNIVA 1260

QY      1261 VERLKEVSETEKAPMOIOETAPSSMPQVGRVEPNRYCLARYEDLOFVLRHINVTNGG 1320
DB      1261 VERLKEVSETEKAPMOIOETAPSSMPQVGRVEPNRYCLARYEDLOFVLRHINVTNGG 1320

QY      1321 EKVIGVGRTAGKGSSTLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPQEPVLF 1380
DB      1321 EKVIGVGRTAGKGSSTLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPQEPVLF 1380

QY      1381 SCSLRNMLDPFSQYSDBEWTSLLELAHKDFVSLPDKLDHECAEGEENISVQORQVCL 1440
DB      1381 SCSLRNMLDPFSQYSDBEWTSLLELAHKDFVSLPDKLDHECAEGEENISVQORQVCL 1440

QY      1441 ARLALRKTILVLDENLTAANDLETTDLIOSTIRFOEDCTVLIARLNTIMYTRIVYL 1500
DB      1441 ARLALRKTILVLDENLTAANDLETTDLIOSTIRFOEDCTVLIARLNTIMYTRIVYL 1500

QY      1501 DKGEIOEYGAPSDLOQRGLFYSMAQDAGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAE 1560
DB      1501 DKGEIOEYGAPSDLOQRGLFYSMAQDAGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAE 1560

QY      1561 PRYTRERRARFVSKKGNCAVHAKNIREQGRFQDVFTTILVDLKWPHHTLIIFTMSPLCSW 1620
DB      1561 PRYTRERRARFVSKKGNCAVHAKNIREQGRFQDVFTTILVDLKWPHHTLIIFTMSPLCSW 1620

QY      1621 LTFAMVWMLIAFAHGLGAGEGNVPCVMSIHSSSAPLFSIEVOYTIIGRSGMTEBEP 1680
DB      1621 LTFAMVWMLIAFAHGLGAGEGNVPCVMSIHSSSAPLFSIEVOYTIIGRSGMTEBEP 1680

QY      1681 LAILILIVONIYVGLMINAIMLGCIFMKTAQAHRAETLLFSKAVVTLRHRGRCLFMLRVG 1740
DB      1681 LAILILIVONIYVGLMINAIMLGCIFMKTAQAHRAETLLFSKAVVTLRHRGRCLFMLRVG 1740

QY      1741 DLKXSMISATTHMOVVRKTTSPBEGVPLHVOYDIPMENGVGNGIFLVAPLIITYVIDS 1800
DB      1741 DLKXSMISATTHMOVVRKTTSPBEGVPLHVOYDIPMENGVGNGIFLVAPLIITYVIDS 1800

QY      1801 NSPLVDLAPSDLHHHDDLEIIVLGEVVEFTGTTQAFVSYLADELILMKQRPVPIAEBD 1860
DB      1801 NSPLVDLAPSDLHHHDDLEIIVLGEVVEFTGTTQAFVSYLADELILMKQRPVPIAEBD 1860

QY      1861 GRYSVDYSKRKGNITKVPPLCTARQLEBDRS 1891
DB      1861 GRYSVDYSKRKGNITKVPPLCTARQLEBDRS 1891
```

RESULT 5
US-09-939-853A-86
; Sequence 86, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Bugeas et al.
; TITLE OF INVENTION: No. US20040039163A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 86
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-86

Query Match 79.4%; Score 7860; DB 3; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADSDPDLMDNNTWNTSNPDTKCFONTVLVWPCFYLMACFPFYLYLSRH 60
DB 1 MALRGFCSADSDPDLMDNNTWNTSNPDTKCFONTVLVWPCFYLMACFPFYLYLSRH 60
QY 61 DRGIQNTPLNKTITAGFLIMVCMADLFYSFWERSRGITLAPVFLVSPLLGITTLLA 120
DB 61 DRGIQNTPLNKTITAGFLIMVCMADLFYSFWERSRGITLAPVFLVSPLLGITTLLA 120
QY 121 TFLQLRRKRGVSSGIMLTFMLVALVCAALILRSKIMTALKEBAQVDFPDITFYVFS 180
DB 121 TFLQLRRKRGVSSGIMLTFMLVALVCAALILRSKIMTALKEBAQVDFPDITFYVFS 180
QY 181 LLLIQLVLSGSDSPLESETIHDNPPCESSASFLSRTFMWITGLIVRGYRQPLEGSD 240
DB 181 LLLIQLVLSGSDSPLESETIHDNPPCESSASFLSRTFMWITGLIVRGYRQPLEGSD 240
QY 241 LWSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVKVVYSSKDPAPKESSTVDANEVEAL 300
DB 241 LWSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVKVVYSSKDPAPKESSTVDANEVEAL 300
QY 301 IVKSPQKEMWNSLFRULYKTFGPRFLMSFFPKAHDLMFSGPOLKLLIFVNDTAPD 360
DB 301 IVKSPQKEMWNSLFRULYKTFGPRFLMSFFPKAHDLMFSGPOLKLLIFVNDTAPD 360
QY 361 WQGFYFVTLFVTAQTLVLAHQYFHCIFVSGMRITKAVIGAVYRKALVITNSARKSSTV 420
DB 361 WQGFYFVTLFVTAQTLVLAHQYFHCIFVSGMRITKAVIGAVYRKALVITNSARKSSTV 420
QY 421 GEIYVNLMSVDAQRFMDLATYINMISAPLQVILALYLLMLNIGPSVLGAVAVMLVAVN 480
DB 421 GEIYVNLMSVDAQRFMDLATYINMISAPLQVILALYLLMLNIGPSVLGAVAVMLVAVN 480
QY 481 AVAMMKTITVOVAMKSKDNRIKLMNELINGIKYLKLYAMBLAFPDKYLALROBELKYLK 540
DB 481 AVAMMKTITVOVAMKSKDNRIKLMNELINGIKYLKLYAMBLAFPDKYLALROBELKYLK 540
QY 541 KSAIYLSAVGTFTWCTPFLVALCTFAVYVITDENNIIDAQTAFYSLAENLIRPINTLP 600
DB 541 KSAIYLSAVGTFTWCTPFLVALCTFAVYVITDENNIIDAQTAFYSLAENLIRPINTLP 600
QY 601 MVISSIVQASVSLRLRILFISHELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660

DB 601 MVISSIVQASVSLRLRILFISHELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660
QY 661 LINGITFSPGALVAVVGVCCKSSLSLALAMDMKVEGHVAIKGSVAIVPQAAWQND 720
DB 661 LINGITFSPGALVAVVGVCCKSSLSLALAMDMKVEGHVAIKGSVAIVPQAAWQND 720
QY 721 SLRENIIFGQLEBPYRSVIOACALLPDEILPSGDRTEIGEXGNVLSGGQKORVSLAR 780
DB 721 SLRENIIFGQLEBPYRSVIOACALLPDEILPSGDRTEIGEXGNVLSGGQKORVSLAR 780
QY 781 AVYSNADIYLPDDPLSAVDHVGKHTFENYIGKGMKAKTRILLVTHSMSTLPVDVITIV 840
DB 781 AVYSNADIYLPDDPLSAVDHVGKHTFENYIGKGMKAKTRILLVTHSMSTLPVDVITIV 840
QY 841 MSGGKISEMSYQBELARDAFAEFLRTYASTEOBDAENGVTVSGPGKEAKOMENG 900
DB 841 MSGGKISEMSYQBELARDAFAEFLRTYASTEOBDAENGVTVSGPGKEAKOMENG 900
QY 901 LVTDAGKQLOROLSSSSSYSGDISRHHNSTABLOKAEAKKEFTWKLMEADKAQTGVKL 960
DB 901 LVTDAGKQLOROLSSSSSYSGDISRHHNSTABLOKAEAKKEFTWKLMEADKAQTGVKL 960
QY 961 SYWMDYKKAIGLPISTLIFLPMCNHYSAASNYLSLWDDPIVNGTOEHTKRLSVYG 1020
DB 961 SYWMDYKKAIGLPISTLIFLPMCNHYSAASNYLSLWDDPIVNGTOEHTKRLSVYG 1020
QY 961 SYWMDYKKAIGLPISTLIFLPMCNHYSAASNYLSLWDDPIVNGTOEHTKRLSVYG 1020
DB 961 SYWMDYKKAIGLPISTLIFLPMCNHYSAASNYLSLWDDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISQIAVFGYSMAVSGIILASRCLAYDLHSLRSFMSFFERPPSGNLVNRFSKEL 1080
DB 1021 ALGISQIAVFGYSMAVSGIILASRCLAYDLHSLRSFMSFFERPPSGNLVNRFSKEL 1080
QY 1081 DTVDMSIPEVIKFMGSLFNVIGACIVILLATPIAIIIPPLGIYFVQRFYVASSROL 1140
DB 1081 DTVDMSIPEVIKFMGSLFNVIGACIVILLATPIAIIIPPLGIYFVQRFYVASSROL 1140
QY 1141 KRLESVRSRSPYSHFNFTLLGVSYTRAFEOERPIHOSDLKVDENQAAVYPSIYANWLA 1200
DB 1141 KRLESVRSRSPYSHFNFTLLGVSYTRAFEOERPIHOSDLKVDENQAAVYPSIYANWLA 1200
QY 1201 VRLCEVNCIYLPALPAVVISRHSLSAGLVGSYSYLOVTTYINMLVRSSEMETNIVA 1260
DB 1201 VRLCEVNCIYLPALPAVVISRHSLSAGLVGSYSYLOVTTYINMLVRSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFNNYCLRYRDLDFVLRIHNVITNGG 1320
DB 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFNNYCLRYRDLDFVLRIHNVITNGG 1320
QY 1321 EKVGIVERTGAQSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTTIIPQDVLV 1380
DB 1321 EKVGIVERTGAQSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTTIIPQDVLV 1380
QY 1381 SGLSRMLDPPSOYSDDEWNTSLELAHLKDFVSALPKLHECAGEGENTSVGQROLVCL 1440
DB 1381 SGLSRMLDPPSOYSDDEWNTSLELAHLKDFVSALPKLHECAGEGENTSVGQROLVCL 1440
QY 1441 ARALIRKTKILLVDEATAVVDLEETDILIOSTIRTOPEDCIVLTIAHRLNTIMDTTRIVL 1500
DB 1441 ARALIRKTKILLVDEATAVVDLEETDILIOSTIRTOPEDCIVLTIAHRLNTIMDTTRIVL 1500
QY 1501 DKGEIQEYGA PSDILQORGLPYSMADAGLV 1531
DB 1501 DKGEIQEYGA PSDILQORGLPYSMADAGLV 1531

RESULT 6
US-10-667-891-6
; Sequence 6, Application US/10667891
; Publication No. US20040171024A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, CHARLES W.
; APPLICANT: BREY, PAUL T.
; APPLICANT: HOLM, INGE
; APPLICANT: GRAILLES, MARINE

APPLICANT: RZHETSKY, ANDREY
 TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND
 FILE OF INVENTION: ANOPHELES
 FILE REFERENCE: 03495-0294-0000
 CURRENT APPLICATION NUMBER: US/10/667,891
 PRIOR FILING DATE: 2003-09-23
 PRIOR APPLICATION NUMBER: 60/413,469
 PRIOR FILING DATE: 2002-09-26
 NUMBER OF SEQ ID NOS: 76
 SOFTWARE: PatentIn Ver. 3.2
 SEQ ID NO: 6
 LENGTH: 1531
 TYPE: PRP
 ORGANISM: Homo sapiens
 US-10-667-891-6

Query Match 79.4%; Score 7860; DB 4; Length 1531;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRFGSADSDPLMDMNTWNTSNPDFTKCFONTVLVWPCYYLMACFPFYLYSRH 60
 DB 1 MALRFGSADSDPLMDMNTWNTSNPDFTKCFONTVLVWPCYYLMACFPFYLYSRH 60
 QY 61 DRGYIOMTPLNKTALGFLIMVCMADLFYSFWEBSRGIFLAVVPLVSPFLGITTLLA 120
 DB 61 DRGYIOMTPLNKTALGFLIMVCMADLFYSFWEBSRGIFLAVVPLVSPFLGITTLLA 120
 QY 121 TFLIOLERRKGVSGGIMLTFMVALVCAALILRSKIMTALKEDAOVDFRDIIFYVFS 180
 DB 121 TFLIOLERRKGVSGGIMLTFMVALVCAALILRSKIMTALKEDAOVDFRDIIFYVFS 180
 QY 181 LLLIQLVLSGSDSPFSETIHDPNCPRESSASFSLRITFTWITGLIVRGYRPLEGSD 240
 DB 181 LLLIQLVLSGSDSPFSETIHDPNCPRESSASFSLRITFTWITGLIVRGYRPLEGSD 240
 QY 241 LMSLNKEDTSEOVVPLVVKWKKSCAKTRKOPVAVVSSKDPAPKESKVDANEVEAL 300
 DB 241 LMSLNKEDTSEOVVPLVVKWKKSCAKTRKOPVAVVSSKDPAPKESKVDANEVEAL 300
 QY 301 IVKSPQKEMNPSLFKVLTKTFPGYFLMSFFPKAIDHLMFSGPOLIKLIFVNDTRAPD 360
 DB 301 IVKSPQKEMNPSLFKVLTKTFPGYFLMSFFPKAIDHLMFSGPOLIKLIFVNDTRAPD 360
 QY 361 WOGFFYVLLFVTAQLOTLVHQYFHCIFVSGMRKTAIVGAVYRKALVITNSARKSTV 420
 DB 361 WOGFFYVLLFVTAQLOTLVHQYFHCIFVSGMRKTAIVGAVYRKALVITNSARKSTV 420
 QY 421 GEIYNLMSVDAOREMDLATYINMTWSAPLOVITLALYLLMLNGPSVLAVVWVLTAMPVN 480
 DB 421 GEIYNLMSVDAOREMDLATYINMTWSAPLOVITLALYLLMLNGPSVLAVVWVLTAMPVN 480
 QY 481 AVNAAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAMELAFKDVLAIRBELKVLK 540
 DB 481 AVNAAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAMELAFKDVLAIRBELKVLK 540
 QY 541 KSAVLAVGFTFWCTPFLVALCTFAVYVITIDENNIIDAQTAEVSALFNILRPLNLP 600
 DB 541 KSAVLAVGFTFWCTPFLVALCTFAVYVITIDENNIIDAQTAEVSALFNILRPLNLP 600
 QY 601 MVISSIVQASVSLKRLIFLSHELEPDSIBRPVKGGGNGNSTVNNATFTARSBPPT 660
 DB 601 MVISSIVQASVSLKRLIFLSHELEPDSIBRPVKGGGNGNSTVNNATFTARSBPPT 660
 QY 661 LMGITFSPGALVAVVGVGCGSKSLLSALLAEMDKVEGHVAIKGSVAVVPQAWIOND 720
 DB 661 LMGITFSPGALVAVVGVGCGSKSLLSALLAEMDKVEGHVAIKGSVAVVPQAWIOND 720
 QY 721 SLRENIIFGQLEPPYRSVIOACALLPDLIELPSGRTEIGEGVNLSSGQKORVSLAR 780
 DB 721 SLRENIIFGQLEPPYRSVIOACALLPDLIELPSGRTEIGEGVNLSSGQKORVSLAR 780
 QY 781 AVYSNADIYFDPLSLAVDAHVGHKIFENVIQPKGMLKNKTRILVTHSMVSLPQVDYIIV 840

DB 781 AVYSNADIYFDPLSLAVDAHVGHKIFENVIQPKGMLKNKTRILVTHSMVSLPQVDYIIV 840
 QY 841 MSGGKISEMSYQOELARDGAFELFTYASTEOEODAEENGVTGSGPKCAKQOMENG 900
 DB 841 MSGGKISEMSYQOELARDGAFELFTYASTEOEODAEENGVTGSGPKCAKQOMENG 900
 QY 901 LMYDSACKOJLOROLSSSSSGDISRHHNSTAEIQXEAKEEFWTKMEADKATQGYKL 960
 DB 901 LMYDSACKOJLOROLSSSSSGDISRHHNSTAEIQXEAKEEFWTKMEADKATQGYKL 960
 QY 961 SVYWDYKAIQGLFISFISFLFMCNHSALASNYMLSMTDDPIVNGTOEHTKRLSYVG 1020
 DB 961 SVYWDYKAIQGLFISFISFLFMCNHSALASNYMLSMTDDPIVNGTOEHTKRLSYVG 1020
 QY 1021 ALGISQGIAYGVGMAVSGIILASRCLHVDLHSLRSFMSFERTPSGNLVNRFSEL 1080
 DB 1021 ALGISQGIAYGVGMAVSGIILASRCLHVDLHSLRSFMSFERTPSGNLVNRFSEL 1080
 QY 1081 DTYDSMTPEVITKMGSLFNVIGACIYIILATPIAIIIPPLGIYFFVQRFYVASSROL 1140
 DB 1081 DTYDSMTPEVITKMGSLFNVIGACIYIILATPIAIIIPPLGIYFFVQRFYVASSROL 1140
 QY 1141 KRLESVSRSPVSHFNFTLGVSVYIRAFEEOEERFIHOSDLKVDENQKAYPSIVANRMLA 1200
 DB 1141 KRLESVSRSPVSHFNFTLGVSVYIRAFEEOEERFIHOSDLKVDENQKAYPSIVANRMLA 1200
 QY 1201 VRLCVCNCIYLPALFAVIRSRHLSAGLVGSYSLOVTTYLNMVLRMSSEMETNIVA 1260
 DB 1201 VRLCVCNCIYLPALFAVIRSRHLSAGLVGSYSLOVTTYLNMVLRMSSEMETNIVA 1260
 QY 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVFNRYCIRYREDLDFVLRHINVTINGG 1320
 DB 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVFNRYCIRYREDLDFVLRHINVTINGG 1320
 QY 1321 EKVGIWGTAGKSLTLGLFRINESABEGIIIDGINAKIGLHDAEFKTTIIPQDPVLF 1380
 DB 1321 EKVGIWGTAGKSLTLGLFRINESABEGIIIDGINAKIGLHDAEFKTTIIPQDPVLF 1380
 QY 1381 SGLSRNMLDPPSOYSDEEVTSLHLAKDFVSLPDKLHCEAGGENTSVGQRQLVCL 1440
 DB 1381 SGLSRNMLDPPSOYSDEEVTSLHLAKDFVSLPDKLHCEAGGENTSVGQRQLVCL 1440
 QY 1441 ARALLRKTILVDEATAVNDLETDDLIOSTIRTOFEDCTVLIARHLANTIMDYTRIVL 1500
 DB 1441 ARALLRKTILVDEATAVNDLETDDLIOSTIRTOFEDCTVLIARHLANTIMDYTRIVL 1500
 QY 1501 DKGEIOEYGAPSDLOORGFLFYSMAKDAGLV 1531
 DB 1501 DKGEIOEYGAPSDLOORGFLFYSMAKDAGLV 1531

RESULT 7
 US-10-807-466-6
 ; Sequence 6, Application US/10807466
 ; Publication No. US20040244066A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ROTH, CHARLES W.
 ; APPLICANT: BREY, PAUL T.
 ; APPLICANT: HOLM, INGE
 ; APPLICANT: GRALLIES, MARINE
 ; APPLICANT: RZHETSKY, ANDREY
 ; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND
 ; FILE REFERENCE: 03495-0294-01000
 ; CURRENT APPLICATION NUMBER: US/10/807,466
 ; CURRENT FILING DATE: 2004-03-24
 ; PRIOR APPLICATION NUMBER: 10/667,891
 ; PRIOR FILING DATE: 2003-09-23
 ; PRIOR APPLICATION NUMBER: 60/413,469
 ; PRIOR FILING DATE: 2002-09-26
 ; NUMBER OF SEQ ID NOS: 139
 ; SOFTWARE: PatentIn Ver. 3.2

```
; SEQ ID NO 6
; LENGTH: 1531
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-807-466-6

Query Match      79.4%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MALRGFCAGDSDPLMDNNTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFYFLYLSRH 60
Db      1 MALRGFCAGDSDPLMDNNTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFYFLYLSRH 60
Qy      61 DRGYIOMTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPILLGITTTLLA 120
Db      61 DRGYIOMTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPILLGITTTLLA 120
Qy      121 TFLIOLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAOVDLFRDITFYVFS 180
Db      121 TFLIOLERRKGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAOVDLFRDITFYVFS 180
Qy      181 LLLIQLVUSCFSDRSPLFSETHIDNPCESSASFLSITFWWINGLIVRGYROPLEGGSD 240
Db      181 LLLIQLVUSCFSDRSPLFSETHIDNPCESSASFLSITFWWINGLIVRGYROPLEGGSD 240
Qy      241 LMSINKEDTSQOVVPLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKYDANEVEAL 300
Db      241 LMSINKEDTSQOVVPLVKNWKECAKTRKQPVKVYSSKDPAPQKSSKYDANEVEAL 300
Qy      301 IVKSPQKMNPSLFLVLYKTGPYFLMSPFFKAIHDLMMFSGPOLIKLIKRVNDTKAPD 360
Db      301 IVKSPQKMNPSLFLVLYKTGPYFLMSPFFKAIHDLMMFSGPOLIKLIKRVNDTKAPD 360
Qy      361 MOGYEYTVLLEFYTACLOTLYLHOYFHCIVSGMRIKTAVIGAVYRKALVITNSAKKSTV 420
Db      361 MOGYEYTVLLEFYTACLOTLYLHOYFHCIVSGMRIKTAVIGAVYRKALVITNSAKKSTV 420
Qy      421 GEIYVNLMSVDAQRFNDLATYINMINSAPLOVTLALYLLMLNIGPSVLGAVAVMLMVEVN 480
Db      421 GEIYVNLMSVDAQRFNDLATYINMINSAPLOVTLALYLLMLNIGPSVLGAVAVMLMVEVN 480
Qy      481 AVAMAMKTITTYOVAMHKSNDNRKIMNEILINGIKVLTLYAMELAFKQVLAIRQEBLKATLK 540
Db      481 AVAMAMKTITTYOVAMHKSNDNRKIMNEILINGIKVLTLYAMELAFKQVLAIRQEBLKATLK 540
Qy      541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTDENNILLDAQTFVSLAFNIRFPLNITLP 600
Db      541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTDENNILLDAQTFVSLAFNIRFPLNITLP 600
Qy      601 MVISSIVQASVSLKRLIFLSHEBELPDSIERRPVKDGGGINSITVRNATFTWASDPT 660
Db      601 MVISSIVQASVSLKRLIFLSHEBELPDSIERRPVKDGGGINSITVRNATFTWASDPT 660
Qy      661 LINGITFSIPEGALVAVVQGVCGKSSLSALLAEMDKYEGHVAITGSVAAYVPOQAMIOND 720
Db      661 LINGITFSIPEGALVAVVQGVCGKSSLSALLAEMDKYEGHVAITGSVAAYVPOQAMIOND 720
Qy      721 SLRENILGCGOLEEYRYSVIOACALLPDLITLPSGDPTIEGKVNLSGGCKQKVSILAR 780
Db      721 SLRENILGCGOLEEYRYSVIOACALLPDLITLPSGDPTIEGKVNLSGGCKQKVSILAR 780
Qy      781 AVYSNADIYLPDDPLSAVDAAHVGKHI FENVIGPKMKLKNKTRILVTSHSASYLPQVDVIV 840
Db      781 AVYSNADIYLPDDPLSAVDAAHVGKHI FENVIGPKMKLKNKTRILVTSHSASYLPQVDVIV 840
Qy      841 MSGKISIMSGSYOELLARDGAFAELKRYASTEOQDAEENGVTGVSQGEKAKOMENGM 900
Db      841 MSGKISIMSGSYOELLARDGAFAELKRYASTEOQDAEENGVTGVSQGEKAKOMENGM 900
Qy      901 LVTTSAGLOLOROUSSSSSYSGDISRHHNSTALEOKAEKKKEETKLMADAAQOGVYL 960
Db      901 LVTTSAGLOLOROUSSSSSYSGDISRHHNSTALEOKAEKKKEETKLMADAAQOGVYL 960
```

```
Qy      961 SVYWDYMKAIQGLFISFLSIFLMCNHVSALASNYWLSLMTDDPIVNGTQEHTKVRLSVYG 1020
Db      961 SVYWDYMKAIQGLFISFLSIFLMCNHVSALASNYWLSLMTDDPIVNGTQEHTKVRLSVYG 1020
Qy      1021 ALGISOGIIVRGYMAVNSIGIILASRCLHYDLHLSIIRSPMSFFERTPSGNLVNRFSEKL 1080
Db      1021 ALGISOGIIVRGYMAVNSIGIILASRCLHYDLHLSIIRSPMSFFERTPSGNLVNRFSEKL 1080
Qy      1081 DTVDSMTPEVIXMKMGSLFNVIACIYILLATPIAAIITPILGIFYFVQGFYVASSRQL 1140
Db      1081 DTVDSMTPEVIXMKMGSLFNVIACIYILLATPIAAIITPILGIFYFVQGFYVASSRQL 1140
Qy      1141 KRLSVSRSPYSHPNETLLGVSVIARAPEQGRPIHOSDLKVDENQKAYPSIVANRWLA 1200
Db      1141 KRLSVSRSPYSHPNETLLGVSVIARAPEQGRPIHOSDLKVDENQKAYPSIVANRWLA 1200
Qy      1201 VRLCEVGNCTVLPAALPAVVISRHSLSAGLVGLSVYSILOVTTYLNMVLRSSSEMETNIVA 1260
Db      1201 VRLCEVGNCTVLPAALPAVVISRHSLSAGLVGLSVYSILOVTTYLNMVLRSSSEMETNIVA 1260
Qy      1261 VERLEKEYSETEKEAPMOIQETAPPSWPOVGRVFEFRNYCLYREDDLFLVLRHINVTINGG 1320
Db      1261 VERLEKEYSETEKEAPMOIQETAPPSWPOVGRVFEFRNYCLYREDDLFLVLRHINVTINGG 1320
Qy      1321 EKVGIVGRTGAGKSLTGLFRINESABGEIITDGINIAKTGLHDLRFKTIITIPQDPVLF 1380
Db      1321 EKVGIVGRTGAGKSLTGLFRINESABGEIITDGINIAKTGLHDLRFKTIITIPQDPVLF 1380
Qy      1381 SGSLRMLNDPSSQVSDERWMTSLHLAKDPVSLPDLDBECABEGGNSVSGOROLVCL 1440
Db      1381 SGSLRMLNDPSSQVSDERWMTSLHLAKDPVSLPDLDBECABEGGNSVSGOROLVCL 1440
Qy      1441 ARALLRKTILVLEAATAVVDLETDLLIOSTIRFOFEDCYLTIJAHRLNTIMDYTRVIVL 1500
Db      1441 ARALLRKTILVLEAATAVVDLETDLLIOSTIRFOFEDCYLTIJAHRLNTIMDYTRVIVL 1500
Qy      1501 DKGEIOEYGAPSDLLQORGLFYSNAXDAGLV 1531
Db      1501 DKGEIOEYGAPSDLLQORGLFYSNAXDAGLV 1531

RESULT 8
US-10-484-577-678
; Sequence 678, Application US/10484577
; Publication No. US20050032724A1
; GENERAL INFORMATION:
; APPLICANT: EPIDAUROS Biotechnologie Aktiengesellschaft
; TITLE OF INVENTION: Means and methods for improved treatment of cancer based on UGT1A
; FILE REFERENCE: P2285PCT-1
; CURRENT APPLICATION NUMBER: US/10/484, 577
; PRIOR FILING DATE: 2004-01-22
; PRIOR APPLICATION NUMBER: PCT/EP 02/08220
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: EP 01 11 7608.8
; PRIOR FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: EP 02011710.7
; NUMBER OF SEQ ID NOS: 683
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 678
; LENGTH: 1531
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-484-577-678

Query Match      79.4%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MALRGFCAGDSDPLMDNNTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFYFLYLSRH 60
Db      1 MALRGFCAGDSDPLMDNNTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFYFLYLSRH 60
```

QY 61 DRGIOMTPLNKTKTALGFLMTIWCADLFYSFMRSGIFLAPVFLVSPTLGITTLA 120
DB 61 DRGIOMTPLNKTKTALGFLMTIWCADLFYSFMRSGIFLAPVFLVSPTLGITTLA 120
QY 121 TFLIOERKGVSSGIMLTFMLVALVICALILRSKIMTALKEDAQVDLFRDIFYYVFS 180
DB 121 TFLIOERKGVSSGIMLTFMLVALVICALILRSKIMTALKEDAQVDLFRDIFYYVFS 180
QY 181 LLLIOVLSCFSDSPLESETIHDNPNCPSSASLSTITWTMTGLIVRGYROPLESD 240
DB 181 LLLIOVLSCFSDSPLESETIHDNPNCPSSASLSTITWTMTGLIVRGYROPLESD 240
QY 241 LMSINKEDTSBOQVPLVKNMKKECAKTRKQPVKVYSSKDPAPKSSKVDANEVEAL 300
DB 241 LMSINKEDTSBOQVPLVKNMKKECAKTRKQPVKVYSSKDPAPKSSKVDANEVEAL 300
QY 301 IVKSPKEMWNSLFVKYKTEGPFYFLMSFFKAIHDLMPFSGPOLKLLIKFVNDTKAPD 360
DB 301 IVKSPKEMWNSLFVKYKTEGPFYFLMSFFKAIHDLMPFSGPOLKLLIKFVNDTKAPD 360
QY 361 WQGFYTVLFTVACLQTLVHOYFHICFVSGMRKTRAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYTVLFTVACLQTLVHOYFHICFVSGMRKTRAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDAQRPMDLATYINMIMSAPOVITLALYLLMLGSPVLAVAVMTLMPVN 480
DB 421 GEIVNLMSVDAQRPMDLATYINMIMSAPOVITLALYLLMLGSPVLAVAVMTLMPVN 480
QY 481 AVMAKTKTYOVAMHMSKDNRIKLMNEILNGIKYLKIYAMELAFKDKVLAIROBELKYLK 540
DB 481 AVMAKTKTYOVAMHMSKDNRIKLMNEILNGIKYLKIYAMELAFKDKVLAIROBELKYLK 540
QY 541 KSAVLAVGFTWCTPPLVALCTPFAVYVITDENNIIDAQTAFLVSLAFNILRPLNLP 600
DB 541 KSAVLAVGFTWCTPPLVALCTPFAVYVITDENNIIDAQTAFLVSLAFNILRPLNLP 600
QY 601 MVLSIYOASVSLKRLIFLSHELEPDSITERPVKOGGNSITVNAATPTMARSDPT 660
DB 601 MVLSIYOASVSLKRLIFLSHELEPDSITERPVKOGGNSITVNAATPTMARSDPT 660
QY 661 LMGTFPSIPGALVAVVGVCGSGSLLSALLAEMDKVEGHVAKGSVAVYPOQAMIOND 720
DB 661 LMGTFPSIPGALVAVVGVCGSGSLLSALLAEMDKVEGHVAKGSVAVYPOQAMIOND 720
QY 721 SLRNNIIFGCOLBEPYRSVYQACALLPDLFILPSGDRTEIGEGVNLSSGQKORSILAR 780
DB 721 SLRNNIIFGCOLBEPYRSVYQACALLPDLFILPSGDRTEIGEGVNLSSGQKORSILAR 780
QY 781 AVVSNADIYLFDDPLSAVDAHVGHIFENVIGPKGMLKNKTRILIVTMSMSTLPQVDYIV 840
DB 781 AVVSNADIYLFDDPLSAVDAHVGHIFENVIGPKGMLKNKTRILIVTMSMSTLPQVDYIV 840
QY 841 MSGGKISEMGSYOELIARDGAFAEFLRTYASTEOBDAEENGVTGSGPGEAKOMENG 900
DB 841 MSGGKISEMGSYOELIARDGAFAEFLRTYASTEOBDAEENGVTGSGPGEAKOMENG 900
QY 901 LVTDSAGKOLOROLSSSSSSSGDISRRHNSIAELQKAEAKKEBTWKLMEADKAQTGVKL 960
DB 901 LVTDSAGKOLOROLSSSSSSSGDISRRHNSIAELQKAEAKKEBTWKLMEADKAQTGVKL 960
QY 961 SYVWDVMAKATGLFISELSTFLPMCNHVSALASNTWLSLMDPDPVNGTOEHTKRLSVYG 1020
DB 961 SYVWDVMAKATGLFISELSTFLPMCNHVSALASNTWLSLMDPDPVNGTOEHTKRLSVYG 1020
QY 1021 ALGISOGIAVFGVMAVSIIGIILASRCLHVDLHSLIRSPMSFEERTPSGULVVRFSKEL 1080
DB 1021 ALGISOGIAVFGVMAVSIIGIILASRCLHVDLHSLIRSPMSFEERTPSGULVVRFSKEL 1080
QY 1081 DTUDSMIPEVYKMGMSLFWNYGACIVITLATAIIPPLGLIYFFVQAFYASSROL 1140
DB 1081 DTUDSMIPEVYKMGMSLFWNYGACIVITLATAIIPPLGLIYFFVQAFYASSROL 1140

QY 1141 KRLSVSRSPVSHFNELLGVSVIRAFEBQERFIHOSDLKVDENOKAYPYSIVANRWLA 1200
DB 1141 KRLSVSRSPVSHFNELLGVSVIRAFEBQERFIHOSDLKVDENOKAYPYSIVANRWLA 1200
QY 1201 VRLBCVNCIYLFALFAVIRSRHLSAGVLGSYSLSQVTTYLNLVWMSSEMETNIVA 1260
DB 1201 VRLBCVNCIYLFALFAVIRSRHLSAGVLGSYSLSQVTTYLNLVWMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMIOETAPPPSWPOVGRVFRNYCLRYRDLDFVLRHINVTING 1320
DB 1261 VERLKEYSETEKAPMIOETAPPPSWPOVGRVFRNYCLRYRDLDFVLRHINVTING 1320
QY 1321 EKVGIVGRTGAGKSLTLGFRINESAGEIIDIINIAKIGLHDLRPFKTIIFQDPVLF 1380
DB 1321 EKVGIVGRTGAGKSLTLGFRINESAGEIIDIINIAKIGLHDLRPFKTIIFQDPVLF 1380
QY 1381 SGSLRMLDPPSOYSDEEWTSLBLAHLKDFVSLPDKLDHECAEGENLSVGOROLVCL 1440
DB 1381 SGSLRMLDPPSOYSDEEWTSLBLAHLKDFVSLPDKLDHECAEGENLSVGOROLVCL 1440
QY 1441 ARALRRTKTLVDEATAVADLETDLIQSTIRQFEDCTVLTIAHRLNTIMDYTRIVL 1500
DB 1441 ARALRRTKTLVDEATAVADLETDLIQSTIRQFEDCTVLTIAHRLNTIMDYTRIVL 1500
QY 1501 DKGEIOEGAPSDLLQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEGAPSDLLQORGLFYSMAKDAGLV 1531

RESULT 9
US-10-889-503-19
Sequence 19, Application US/10889503
Publication No. US20050063968A1
GENERAL INFORMATION:
APPLICANT: Fox Chase Cancer Center
APPLICANT: Kruth, Gary D.
APPLICANT: Lee, Kun
APPLICANT: Belinsky, Martin G.
TITLE OF INVENTION: MRP-Related ABC Transporter Encoding
FILE REFERENCE: FCCC 98-02
CURRENT APPLICATION NUMBER: US/10/889,503
PRIOR FILING DATE: 2004-07-12
PRIOR APPLICATION NUMBER: US/09/647,140
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US99/06644
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 60/079,759
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/095,153
PRIOR FILING DATE: 1998-08-03
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 19
LENGTH: 1531
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-889-503-19
Query Match 79.4%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCASDSDPLMDMNTYNTSNPDPFKCFQNTLVVWPCFYLMACPPFYFLYSRH 60
DB 1 MALRGFCASDSDPLMDMNTYNTSNPDPFKCFQNTLVVWPCFYLMACPPFYFLYSRH 60
QY 61 DRGIOMTPLNKTKTALGFLMTIWCADLFYSFMRSGIFLAPVFLVSPTLGITTLA 120
DB 61 DRGIOMTPLNKTKTALGFLMTIWCADLFYSFMRSGIFLAPVFLVSPTLGITTLA 120
QY 121 TFLIOERKGVSSGIMLTFMLVALVICALILRSKIMTALKEDAQVDLFRDITFYVFS 180
DB 121 TFLIOERKGVSSGIMLTFMLVALVICALILRSKIMTALKEDAQVDLFRDITFYVFS 180

```

121 TFLIQLERRKVGSSGIMLTFMLVALVCALILRSKIMTALKEBAQVDLFFDITFYVFS 180
181 LLLIQLVLSGSDSPLEFSETIHDNPNCPRESSASFLSRTFMWITGLIVRGYRQPLEGSD 240
181 LLLIQLVLSGSDSPLEFSETIHDNPNCPRESSASFLSRTFMWITGLIVRGYRQPLEGSD 240
241 LMSLINKEDTSQVVPVLYVKNMKECAKTRKQPVKVYVSSKDPAPKSSKVDANEVEAL 300
241 LMSLINKEDTSQVVPVLYVKNMKECAKTRKQPVKVYVSSKDPAPKSSKVDANEVEAL 300
301 IVKSPQKEMNSLFRKVLKTFGPFYFMSFFKAIHDLMMFSGPOLIKLLIFVNDTKAPD 360
301 IVKSPQKEMNSLFRKVLKTFGPFYFMSFFKAIHDLMMFSGPOLIKLLIFVNDTKAPD 360
361 WQGFYFVLLFVTAQLOTLVLIHQYFHI CFVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420
361 WQGFYFVLLFVTAQLOTLVLIHQYFHI CFVSGMRKTAIVIGAVYRKALVITNSARKSSTV 420
421 GEIIVNLSVDAQRFDMLATYINMISAPLOVILALYLLMLNGPSVLGAVAVMYLMEVFN 480
421 GEIIVNLSVDAQRFDMLATYINMISAPLOVILALYLLMLNGPSVLGAVAVMYLMEVFN 480
481 AVAMAKTKTYOVAMHKSNDRIKLMNELLNGIKYKLYAMELAFKDKVLAIRQEBLKVK 540
481 AVAMAKTKTYOVAMHKSNDRIKLMNELLNGIKYKLYAMELAFKDKVLAIRQEBLKVK 540
541 KSAIYASAVGTFTWVCTPFLVALCTFAVAVTIDENNILDAQTAFAVSALFNILRPPLNLP 600
541 KSAIYASAVGTFTWVCTPFLVALCTFAVAVTIDENNILDAQTAFAVSALFNILRPPLNLP 600
601 MVISSIVQASVSLKRLRIFISHEBELPDSIERRPVKDGGGINSITVRNATFTWASDPT 660
601 MVISSIVQASVSLKRLRIFISHEBELPDSIERRPVKDGGGINSITVRNATFTWASDPT 660
661 LMGITFSPREGALVAVVQVCGGKSSLLSALLAENDKXEGHVALIGSAVYVPOQAMIOND 720
661 LMGITFSPREGALVAVVQVCGGKSSLLSALLAENDKXEGHVALIGSAVYVPOQAMIOND 720
721 SLRENIIFGCOLEBPYRSVIOACALPDLLEILPSGDTIEIGEKVNLSSGOKORVSLAR 780
721 SLRENIIFGCOLEBPYRSVIOACALPDLLEILPSGDTIEIGEKVNLSSGOKORVSLAR 780
781 AVYSNADITLFPDDPLSAVDAHVGKHFENVI GPKMKNKTRILVTHSMYSLPOVDVIV 840
781 AVYSNADITLFPDDPLSAVDAHVGKHFENVI GPKMKNKTRILVTHSMYSLPOVDVIV 840
841 MSGGKISMGSYOELIARDGAFAEFLRTYASTEOQDAEENGVTGVSQPKAKEMNGM 900
841 MSGGKISMGSYOELIARDGAFAEFLRTYASTEOQDAEENGVTGVSQPKAKEMNGM 900
901 LVTSAGAGLOROLSSSSSSSGDLSRHNSSTAELQKAEKKEETKLMBADKAOTGOYKL 960
901 LVTSAGAGLOROLSSSSSSSGDLSRHNSSTAELQKAEKKEETKLMBADKAOTGOYKL 960
961 SVYNDYMAKIGLFIISFLSIFLPMCNHVSALASNYMLSMTDDPIVNGTOEHTKVALSYG 1020
961 SVYNDYMAKIGLFIISFLSIFLPMCNHVSALASNYMLSMTDDPIVNGTOEHTKVALSYG 1020
1021 ALGISOGIAVPGYSMAVSIIGGILASRCIHLVLIHSILRS PMSFFERTSGNUNRFSKEL 1080
1021 ALGISOGIAVPGYSMAVSIIGGILASRCIHLVLIHSILRS PMSFFERTSGNUNRFSKEL 1080
1081 DTVOSMIEVIMKMFSGLPNVIGACIVILLATPIAIIIPPLGLYFFVQRYVASSROL 1140
1081 DTVOSMIEVIMKMFSGLPNVIGACIVILLATPIAIIIPPLGLYFFVQRYVASSROL 1140
1141 KRLSESVRSPPYSHNETLIGSVIRAREEOERFIHOSDLKVDENOKAYPYSIVANRMLA 1200
1141 KRLSESVRSPPYSHNETLIGSVIRAREEOERFIHOSDLKVDENOKAYPYSIVANRMLA 1200
1201 VRLCEVNCIYVPAALFAVISHSLSAGLVGIVSVSYSLQVTTYINMLVRMSSEMETNIVA 1260
1201 VRLCEVNCIYVPAALFAVISHSLSAGLVGIVSVSYSLQVTTYINMLVRMSSEMETNIVA 1260

```

```

1201 VRLCEVNCIYVPAALFAVISHSLSAGLVGIVSVSYSLQVTTYINMLVRMSSEMETNIVA 1260
1261 VERLKEYSETEKBAQWIOETAPSSWPQYGRVEFRNYCLRYREDLFEVLRHINVTINGG 1320
1261 VERLKEYSETEKBAQWIOETAPSSWPQYGRVEFRNYCLRYREDLFEVLRHINVTINGG 1320
1321 EKVGIVERTGAGKSSLTGLFPRINESAGEIIIDGINIAKIGLHDLAFKTIITIPQDPVLF 1380
1321 EKVGIVERTGAGKSSLTGLFPRINESAGEIIIDGINIAKIGLHDLAFKTIITIPQDPVLF 1380
1381 SGLSRMLDPSQVSDDEWNTSLAHLKDFVSALPDKLHECAEGGENLSVGOROLVCL 1440
1381 SGLSRMLDPSQVSDDEWNTSLAHLKDFVSALPDKLHECAEGGENLSVGOROLVCL 1440
1441 ARALIRKTIILVDEATAVDELTDLIQSTIRTOFEDCTVLIATIRLANTIMDYTRYVL 1500
1441 ARALIRKTIILVDEATAVDELTDLIQSTIRTOFEDCTVLIATIRLANTIMDYTRYVL 1500
1501 DKGEIOEYGAUSDLLQORGLFYMAKADAGLV 1531
1501 DKGEIOEYGAUSDLLQORGLFYMAKADAGLV 1531

RESULT 10
US-10-756-149-5033
; Sequence 5033, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natesha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: file
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5033
; LENGTH: 1531
; TYPE: PRF
; ORGANISM: Homo Sapiens
US-10-756-149-5033

Query Match 79.4%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db 361 WQGYFTVLLFVTACLOTVLHGYHICFVSGMRKTAIVGAVYRKALVITNSARKSTV 420
Qy 421 GEIYNLMSVDAQRPMDLATYINMTMSAPIOVTLALYLMLNLPSPSLAGVAVMLAMPVN 480
Db 421 GEIYNLMSVDAQRPMDLATYINMTMSAPIOVTLALYLMLNLPSPSLAGVAVMLAMPVN 480
Qy 481 AVMAKTKTYOVAMHMSKDNRIKLMNEILNGIKVLYAMBLAFKDLKAIROBELKVLK 540
Db 481 AVMAKTKTYOVAMHMSKDNRIKLMNEILNGIKVLYAMBLAFKDLKAIROBELKVLK 540
Qy 541 KSAVLSAVGTFTWCTPELVALCTPAVYVITIDENNIIDAQTAFAVSLAFNLRPELNL 600
Db 541 KSAVLSAVGTFTWCTPELVALCTPAVYVITIDENNIIDAQTAFAVSLAFNLRPELNL 600
Qy 601 MVTSSIVQASVSLKRLIFLSHELEPDSIRRPVKOGGNGSTTVNATTAARSPT 660
Db 601 MVTSSIVQASVSLKRLIFLSHELEPDSIRRPVKOGGNGSTTVNATTAARSPT 660
Qy 661 LNCITFSPREGALVAVVGVCGKSSLSALLAEMDKVEGVAIKGSVAVYPOQAMQND 720
Db 661 LNCITFSPREGALVAVVGVCGKSSLSALLAEMDKVEGVAIKGSVAVYPOQAMQND 720
Qy 721 SLRENILFGCQLEBPYRSVIOACALLPDLIELPSGDRTEIGEKVNLSGGQKQVSLAR 780
Db 721 SLRENILFGCQLEBPYRSVIOACALLPDLIELPSGDRTEIGEKVNLSGGQKQVSLAR 780
Qy 781 AVYNAIIVLFDDLSAVDAVGHIFENYIGPKMLKNTKRIIVTMSKTLPOVDYIV 840
Db 781 AVYNAIIVLFDDLSAVDAVGHIFENYIGPKMLKNTKRIIVTMSKTLPOVDYIV 840
Qy 841 MSGKISSEMGSYQELLRDGAFAEFLRTYASTEOBDAEENGVTGSGPGKEAKQMNMG 900
Db 841 MSGKISSEMGSYQELLRDGAFAEFLRTYASTEOBDAEENGVTGSGPGKEAKQMNMG 900
Qy 901 LVTDSACKOLOROLSSSSSYSGDISRHHNSTAELQKAKKEETWKLMEADKAQTGVKL 960
Db 901 LVTDSACKOLOROLSSSSSYSGDISRHHNSTAELQKAKKEETWKLMEADKAQTGVKL 960
Qy 961 SVYWDYKAIKGLFISFLIFPMCNHVSALASNTWLSMTDDPVTNGTOEHTKRLSVYG 1020
Db 961 SVYWDYKAIKGLFISFLIFPMCNHVSALASNTWLSMTDDPVTNGTOEHTKRLSVYG 1020
Qy 1021 ALGISQIAVFGYSMAVSIQGIILASRCLHVDLHSHIRSPMSFEPRTPSGNLVRFSEKL 1080
Db 1021 ALGISQIAVFGYSMAVSIQGIILASRCLHVDLHSHIRSPMSFEPRTPSGNLVRFSEKL 1080
Qy 1081 DTVDSMIPVYIKMFMGSLFNVIGACIYILATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
Db 1081 DTVDSMIPVYIKMFMGSLFNVIGACIYILATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
Qy 1141 KRLSVRSRPSYSHFNETLLGVSYTRAPEBOERTIHOSDLKVDENQKAYPSIVANWLA 1200
Db 1141 KRLSVRSRPSYSHFNETLLGVSYTRAPEBOERTIHOSDLKVDENQKAYPSIVANWLA 1200
Qy 1201 VRLCEVCNACIVLPALPAVISRHSLSAGLVSVSYLQVTTYINMLVRMSSEMTNIVA 1260
Db 1201 VRLCEVCNACIVLPALPAVISRHSLSAGLVSVSYLQVTTYINMLVRMSSEMTNIVA 1260
Qy 1261 VERLKEVSETEKEKPMOIOETAPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVTNGG 1320
Db 1261 VERLKEVSETEKEKPMOIOETAPSSWPQVGRVEFRNYCLRYREDLDFVLRHINVTNGG 1320
Qy 1321 EKVGIVRTGAGKSSLTGLFRINESABGEIINDGINIAKIGLHDLFFKTTIIPQDPVLF 1380
Db 1321 EKVGIVRTGAGKSSLTGLFRINESABGEIINDGINIAKIGLHDLFFKTTIIPQDPVLF 1380
Qy 1381 SGLSRMLDPPSOYSDBEVMTSLELAHKDFVSLPKLDBECAGGBENISVQROROIYCL 1440
Db 1381 SGLSRMLDPPSOYSDBEVMTSLELAHKDFVSLPKLDBECAGGBENISVQROROIYCL 1440
Qy 1441 ARALLRKTILVLEDEATAVDLETDLIQSTIRTOFEDCTVLTAHRLANTIMDTRIVL 1500
Db 1441 ARALLRKTILVLEDEATAVDLETDLIQSTIRTOFEDCTVLTAHRLANTIMDTRIVL 1500

Qy 1501 DKGEIOEGAPSDLLQORGLFYSMAXDAGV 1531
Db 1501 DKGEIOEGAPSDLLQORGLFYSMAXDAGV 1531

RESULT 11
US-09-939-853A-87
; Sequence 87, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgees et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939, 853A
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 1515
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-87

Query Match 78.5%; Score 7769; DB 3; Length 1515;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 DMVNTWNTSPDPTKCRQNTVLVWVPCFYLMACPPFYFLSHRDGYIOMTPLNKTXTA 76
Db 1 DMVNTWNTSPDPTKCRQNTVLVWVPCFYLMACPPFYFLSHRDGYIOMTPLNKTXTA 60

Qy 77 LGFLMLVWCADLFYSFEMERSRGIPLAPVFLVSPTLIGITLLATPLIQERRKGVQSSG 136
Db 61 LGFLMLVWCADLFYSFEMERSRGIPLAPVFLVSPTLIGITLLATPLIQERRKGVQSSG 120

Qy 137 IMTFMVLVLCALAIRSKIMTALKEDAOVDLFRDITFYVYSSLLIQVLVSCFSRSP 196
Db 121 IMTFMVLVLCALAIRSKIMTALKEDAOVDLFRDITFYVYSSLLIQVLVSCFSRSP 180

Qy 197 LFSETHDPNCPRESSASFLSRITFMWITGLIVRGYRQPLEGSDMLINKEDTSEQVVPV 256
Db 181 LFSETHDPNCPRESSASFLSRITFMWITGLIVRGYRQPLEGSDMLINKEDTSEQVVPV 240

Qy 257 LVKNMKKECAKTRKQPVKVYVSSKDPAPQKSSKVDANBEVALIVKSPKEMNPSLFKV 316
Db 241 LVKNMKKECAKTRKQPVKVYVSSKDPAPQKSSKVDANBEVALIVKSPKEMNPSLFKV 300

Qy 317 LYTFPGPYFLMSPFPAIHDLNMFSGPQIIKLLIKTVNDTKADWDQGYFTVLLFVTACL 376
Db 301 LYTFPGPYFLMSPFPAIHDLNMFSGPQIIKLLIKTVNDTKADWDQGYFTVLLFVTACL 360

Qy 377 QTVLHMQPHICVSGMRKITAIVIGAVYRKALVITNSARKSSVGEIVNLMSVDAQRMD 436
Db 361 QTVLHMQPHICVSGMRKITAIVIGAVYRKALVITNSARKSSVGEIVNLMSVDAQRMD 420

Qy 437 LATYINMISAPQVILALYILMLNLPSPSLAGVAVMLAVPNAVMAKTKTYQVAHMK 496
Db 421 LATYINMISAPQVILALYILMLNLPSPSLAGVAVMLAVPNAVMAKTKTYQVAHMK 480

Qy 497 SKONRIKLMNEILNGIKVLYAMBLAFKDLKAIROBELKVLKSAVYSAVGTFTWCT 556
Db 481 SKONRIKLMNEILNGIKVLYAMBLAFKDLKAIROBELKVLKSAVYSAVGTFTWCT 540

Qy 557 PFLVALCTPAVYVITIDENNIIDAQTAFAVSLAFNLRPELNLPMVTSSIVQASVSLKRL 616

```
|||||
Db      541 PFLVALCTFAVYVITIDENNIIIDAQTAFAVSIALFNILRPPLNILPMVSISSIVQASVSLKRL 600
Qy      617 RIFLSHEBELPDSIERRPVYDGGGNSITVRNATFTMARSPPTLNGITFSIPGALVAV 676
Db      601 RIFLSHEBELPDSIERRPVYDGGGNSITVRNATFTMARSPPTLNGITFSIPGALVAV 660
Qy      677 VGVGCGKSSLLSALLAEMDKVEGHVAKGSVAVYPOQAWIQNDSLRNLIIFGQLEBPY 736
Db      661 VGVGCGKSSLLSALLAEMDKVEGHVAKGSVAVYPOQAWIQNDSLRNLIIFGQLEBPY 720
Qy      737 YRSVIOACALLPDLIELPSGDRTEIGEGKVNLSGGQKORVSLARAVYSNADIYLPDPLS 796
Db      721 YRSVIOACALLPDLIELPSGDRTEIGEGKVNLSGGQKORVSLARAVYSNADIYLPDPLS 780
Qy      797 AYDAHVGHKIFENYIGPKGMLKNKTRILVTHSMSTLPVDVYIYVSGKISBMSYOBLL 856
Db      781 AYDAHVGHKIFENYIGPKGMLKNKTRILVTHSMSTLPVDVYIYVSGKISBMSYOBLL 840
Qy      857 ARDGAFAEFLRTVASTEOBDAEENGVTGVSQPKGEAKOMENGMVLTDSAGKOLQROLSS 916
Db      841 ARDGAFAEFLRTVASTEOBDAEENGVTGVSQPKGEAKOMENGMVLTDSAGKOLQROLSS 900
Qy      917 SSSYSGDISRHHNSTAELOKAEAKKEETWKLMEADKAOTGOYKLSVYDYMKAIGLFTSF 976
Db      901 SSSYSGDISRHHNSTAELOKAEAKKEETWKLMEADKAOTGOYKLSVYDYMKAIGLFTSF 960
Qy      977 LSTIFEMCNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSVYGAALGSOGIAVFGYSMA 1036
Db      961 LSTIFEMCNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSVYGAALGSOGIAVFGYSMA 1020
Qy      1037 VSIIGIILASRCLAHVDLHSLRSFMSFPERTPSGNLVNRFSKELDTVDMSIPEVIKMEWG 1096
Db      1021 VSIIGIILASRCLAHVDLHSLRSFMSFPERTPSGNLVNRFSKELDTVDMSIPEVIKMEWG 1080
Qy      1097 SLENVIGACIYITLATAPIAIIIPPLGIYFFVQAFYVASSRQKLKLESVSRSPYSHFN 1156
Db      1081 SLENVIGACIYITLATAPIAIIIPPLGIYFFVQAFYVASSRQKLKLESVSRSPYSHFN 1140
Qy      1157 ETLGVSIVIRAFEEBOERFIHOSDLKVDENOKAYPSIYANRMLAVRLCEVGNCTYLPAL 1216
Db      1141 ETLGVSIVIRAFEEBOERFIHOSDLKVDENOKAYPSIYANRMLAVRLCEVGNCTYLPAL 1200
Qy      1217 FAVISRHSLSAGVGLSVSYSLOVTTYLNMLVRMSSEMETNI VAVERLKEYSETEKEAPW 1276
Db      1201 FAVISRHSLSAGVGLSVSYSLOVTTYLNMLVRMSSEMETNI VAVERLKEYSETEKEAPW 1260
Qy      1277 QIOETAPSSWPQYGRVFEFRNYCLARYEDLDPVLRHINVTINGEKVGI VERTGAGKSSL 1336
Db      1261 QIOETAPSSWPQYGRVFEFRNYCLARYEDLDPVLRHINVTINGEKVGI VERTGAGKSSL 1320
Qy      1337 TLGIFRIINESAGEIIIDIGINIAKIGHDLRFKTIIPODPYLPFSGSLRMLDPPSOYSD 1396
Db      1321 TLGIFRIINESAGEIIIDIGINIAKIGHDLRFKTIIPODPYLPFSGSLRMLDPPSOYSD 1380
Qy      1397 EEVWTSLELAHKDFVSALPDKLDHECAEGGENLSVGOQOLVCLARALLRKTKIILVDEA 1456
Db      1381 EEVWTSLELAHKDFVSALPDKLDHECAEGGENLSVGOQOLVCLARALLRKTKIILVDEA 1440
Qy      1457 TAAVDELETDLIQSTIRFOEDCTVLTIAHRLNTIMDYTRYIVLDKGEIOEYGAESDILQ 1516
Db      1441 TAAVDELETDLIQSTIRFOEDCTVLTIAHRLNTIMDYTRYIVLDKGEIOEYGAESDILQ 1500
Qy      1517 ORGLFYSMAXKADAGLV 1531
Db      1501 ORGLFYSMAXKADAGLV 1515

RESULT 12
US-10-618-281-42
; Sequence 42, Application US/10618281
; Publication No. US20040219609A1
; GENERAL INFORMATION:
```

```
APPLICANT: Day, Anthony G.
; APPLICANT: Estell, David A.
; APPLICANT: Lyons, Eric H.
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: Methods for Modulating Proteins Not
; TITLE OF INVENTION: Previously known as Proteases
; FILE REFERENCE: GC773-2
; CURRENT APPLICATION NUMBER: US/10/618,281
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: US 60/395,325
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 1515
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-618-281-42

Query Match      78.4%; Score 7763; DB 5; Length 1515;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1514; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      17 DWNTVNTSNDFPKCFQONTYLVWVPCFYLMACPFYFLYSRHRDGYIQMTPLNKTKTA 76
Db      1 DWNTVNTSNDFPKCFQONTYLVWVPCFYLMACPFYFLYSRHRDGYIQMTPLNKTKTA 60
Qy      77 LGFLMLIVCANLDYFSWERSRGIFLAPVFLVSPTLGITTLATFLIOLERRGVQSSG 136
Db      61 LGFLMLIVCANLDYFSWERSRGIFLAPVFLVSPTLGITTLATFLIOLERRGVQSSG 120
Qy      137 IMLTFWVALVICALAIIRSKIMTALKEADAQYDFRDIITFYVYFSLILQVLSCFSRSP 196
Db      121 IMLTFWVALVICALAIIRSKIMTALKEADAQYDFRDIITFYVYFSLILQVLSCFSRSP 180
Qy      197 LFSETHIDPNCPESSASFLSRITFWMITGLIYNGYQPLBGSULMSLNKEDISEQVVPV 256
Db      181 LFSETHIDPNCPESSASFLSRITFWMITGLIYNGYQPLBGSULMSLNKEDISEQVVPV 240
Qy      257 LVKWKKECAKTRQOPKVVYSSKDPQOPKESSVVDNENEVEALIVSPQKEMNPULFKV 316
Db      241 LVKWKKECAKTRQOPKVVYSSKDPQOPKESSVVDNENEVEALIVSPQKEMNPULFKV 300
Qy      317 LYKTFGYFLMSFFFKAIHDLMPFSGPOLIKLIFVNDTKAPDMOGFYTVLLFVTAQL 376
Db      301 LYKTFGYFLMSFFFKAIHDLMPFSGPOLIKLIFVNDTKAPDMOGFYTVLLFVTAQL 360
Qy      377 QTLVLHGYFHI CFVSGMRIKTAVIGAYYRKALVITNSARKSSTVGEIYVNLMSVDAORFMD 436
Db      361 QTLVLHGYFHI CFVSGMRIKTAVIGAYYRKALVITNSARKSSTVGEIYVNLMSVDAORFMD 420
Qy      437 LATYNNIMWSAPLOVITLALVYLMNLGPSYLAGAVWVWVNAWNAWKTKTYQVAHMK 496
Db      421 LATYNNIMWSAPLOVITLALVYLMNLGPSYLAGAVWVWVNAWNAWKTKTYQVAHMK 480
Qy      497 SKDNRKIKMEIILNGIKVILKYAMELAFKQVLAIROBELKVLKKSAYLSAVGTFTWCT 556
Db      481 SKDNRKIKMEIILNGIKVILKYAMELAFKQVLAIROBELKVLKKSAYLSAVGTFTWCT 540
Qy      557 PFLVALCTFAVYVITIDENNIIIDAQTAFAVSIALFNILRPPLNILPMVSISSIVQASVSLKRL 616
Db      541 PFLVALCTFAVYVITIDENNIIIDAQTAFAVSIALFNILRPPLNILPMVSISSIVQASVSLKRL 600
Qy      617 RIFLSHEBELPDSIERRPVYDGGGNSITVRNATFTMARSPPTLNGITFSIPGALVAV 676
Db      601 RIFLSHEBELPDSIERRPVYDGGGNSITVRNATFTMARSPPTLNGITFSIPGALVAV 660
Qy      677 VGVGCGKSSLLSALLAEMDKVEGHVAKGSVAVYPOQAWIQNDSLRNLIIFGQLEBPY 736
Db      661 VGVGCGKSSLLSALLAEMDKVEGHVAKGSVAVYPOQAWIQNDSLRNLIIFGQLEBPY 720
Qy      737 YRSVIOACALLPDLIELPSGDRTEIGEGKVNLSGGQKORVSLARAVYSNADIYLPDPLS 796
```

```

Db 721 YRSVIOACALLPDLLEILPSGDRTEIGEKVNLGGQKORVSLARAVSNADIVLFDDEL 780
Qy 797 AVDAHVGHIFENYIGPKMKLKNKTRILVTHSMYLPQVDVITVMSGGKISEMSYQELL 856
Db 781 AVDAHVGHIFENYIGPKMKLKNKTRILVTHSMYLPQVDVITVMSGGKISEMSYQELL 840
Qy 857 ARDGAFAEFLRTVASTEOBDAENGVTGSGPKGAQOMENGLVTSASAKOLOROLSS 916
Db 841 ARDGAFAEFLRTVASTEOBDAENGVTGSGPKGAQOMENGLVTSASAKOLOROLSS 900
Qy 917 SSSYSGDISRHNSHTAELOKAEAKEETWKLMEADKAQTOGVKLSVYDWYKAIIGLTFISF 976
Db 901 SSSYSGDISRHNSHTAELOKAEAKEETWKLMEADKAQTOGVKLSVYDWYKAIIGLTFISF 960
Qy 977 LSTLFPMCNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSVGALGISOGIAVFGYMA 1036
Db 961 LSTLFPMCNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSVGALGISOGIAVFGYMA 1020
Qy 1037 VSIIGIILASRCLAHVDLHSLIRSFMSPFERTPSGNLVNRFSELDVDSMIPEVIKPMFG 1096
Db 1021 VSIIGIILASRCLAHVDLHSLIRSFMSPFERTPSGNLVNRFSELDVDSMIPEVIKPMFG 1080
Qy 1097 SLFNVIGACIVILLATPIAIIIPPLGLIYFVORFYVASSRODKRLESVRSRPSYSHN 1156
Db 1081 SLFNVIGACIVILLATPIAIIIPPLGLIYFVORFYVASSRODKRLESVRSRPSYSHN 1140
Qy 1157 ETLIGSVIRAFEEQERFIHOSDLKVDENOKAYYPSIVANRWLAVERLECVNCTIVLPAL 1216
Db 1141 ETLIGSVIRAFEEQERFIHOSDLKVDENOKAYYPSIVANRWLAVERLECVNCTIVLPAL 1200
Qy 1217 FAVISRSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETIVAVEBRLKESSETEKRAPW 1276
Db 1201 FAVISRSLSAGLVGLSVSYSLQVTTYLNLVMSSEMETIVAVEBRLKESSETEKRAPW 1260
Qy 1277 QIOETAPSSWPQYGRVFRNYCYLRBEDLDFVLRHINVTINGEKYGIVRTGAKSSL 1336
Db 1261 QIOETAPSSWPQYGRVFRNYCYLRBEDLDFVLRHINVTINGEKYGIVRTGAKSSL 1320
Qy 1337 TLGIFRINESAEGIIIDGINIAKIGLHDLRFKTIITIPDPVLFSSIRMLNDPFSQYSD 1396
Db 1321 TLGIFRINESAEGIIIDGINIAKIGLHDLRFKTIITIPDPVLFSSIRMLNDPFSQYSD 1380
Qy 1397 EEWTSJLELAHLKDFVSLPDKDHECAEGGENSVGOROLVCARALLRKTKLTVLDEA 1456
Db 1381 EEWTSJLELAHLKDFVSLPDKDHECAEGGENSVGOROLVCARALLRKTKLTVLDEA 1440
Qy 1457 TAAVDELTDDLIOSTIRTOEDCTVLTIAHRLNTIMDYTRYIVLDKGEIOEYGAPSDLQ 1516
Db 1441 TAAVDELTDDLIOSTIRTOEDCTVLTIAHRLNTIMDYTRYIVLDKGEIOEYGAPSDLQ 1500
Qy 1517 QRGIFYSMACDAGLV 1531
Db 1501 QRGIFYSMACDAGLV 1515

```

```

RESULT 13
US-10-408-765A-1718
; Sequence 1718, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Bojn D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Marnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408.765A
; NUMBER OF SEQ ID NOS: 3077

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1718
; LENGTH: 1388
; TYPE: FRT
; ORGANISM: Homo sapiens
US-10-408-765A-1718

Query Match      71.1%; Score 7037; DB 4; Length 1388;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1387; Conservative 0; Mismatches 1; Indels 42; Gaps 1;

Qy 67 MTPLNKTKTALGFLMTIVCWADLFYSFMEBSRGIFLAPVPLVSPTLGITTLATPFIQL 126
Db 1 MTPLNKTKTALGFLMTIVCWADLFYSFMEBSRGIFLAPVPLVSPTLGITTLATPFIQL 60
Qy 127 ERRKVOSSGIMLTFWLVALVICALAIIRSKIMTAKEDAQOVLFRDITFYVYFSLLIQL 186
Db 61 ERRKVOSSGIMLTFWLVALVICALAIIRSKIMTAKEDAQOVLFRDITFYVYFSLLIQL 97
Qy 187 VLSCFSRSLPFSFTIHDPNCPRESSASFLSRTFMWITGLIYRGYROPLEGSDMLNK 246
Db 98 -----NCPRESSASFLSRTFMWITGLIYRGYROPLEGSDMLNK 138
Qy 247 EDTSEQVVPVLVKNMKKECAKTRKQPYKVYSSKDPAQPKRSSKVDANEVEALIVKSPQ 306
Db 139 EDTSEQVVPVLVKNMKKECAKTRKQPYKVYSSKDPAQPKRSSKVDANEVEALIVKSPQ 198
Qy 307 KENWPSLFKVLKTFGFPYFLMSFFPKAIDHLMFSGPQIIKLKIFVNDTKAPDMQGYFY 366
Db 199 KENWPSLFKVLKTFGFPYFLMSFFPKAIDHLMFSGPQIIKLKIFVNDTKAPDMQGYFY 258
Qy 367 TULLPFTACIOTVLHGYFHCIFVSGKRITAYIGAYRQALVTTNSARKSSYTGELVNL 426
Db 259 TULLPFTACIOTVLHGYFHCIFVSGKRITAYIGAYRQALVTTNSARKSSYTGELVNL 318
Qy 427 MSYDAORFMDLATYINNIWASAPLOVITATLVLMLNTPSYLAGAVVWLVVAVNAAWK 486
Db 319 MSYDAORFMDLATYINNIWASAPLOVITATLVLMLNTPSYLAGAVVWLVVAVNAAWK 378
Qy 487 TKTYOVAHMSKONRIKLAMEIINGIKVLKLYAMELAFKQKVALIROEBLKVLKKSAYLS 546
Db 379 TKTYOVAHMSKONRIKLAMEIINGIKVLKLYAMELAFKQKVALIROEBLKVLKKSAYLS 438
Qy 547 AVGTFTVCTPRPVYALCTPAVYTTIDENNLIDQAOTAVSLATRNILFPLNLPMTYSSI 606
Db 439 AVGTFTVCTPRPVYALCTPAVYTTIDENNLIDQAOTAVSLATRNILFPLNLPMTYSSI 498
Qy 607 VQASVSLKRLRIFLSHEELPDSTERRPVQDGGTNSITVRNATFTWARSDDPTLNGITF 666
Db 499 VQASVSLKRLRIFLSHEELPDSTERRPVQDGGTNSITVRNATFTWARSDDPTLNGITF 558
Qy 667 SIEGALVAVVGQVCGKSSLSALLAEMDKVEGHAIKGSVAVVPQAMITONDSIRENI 726
Db 559 SIEGALVAVVGQVCGKSSLSALLAEMDKVEGHAIKGSVAVVPQAMITONDSIRENI 618
Qy 727 LFGCOLBEPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLGGQKORVSLARAVYNA 786
Db 619 LFGCOLBEPYRSVIOACALLPDLLEILPSGDRTEIGEKVNLGGQKORVSLARAVYNA 678
Qy 787 DITLPDPLSAVAHAGKHIFENVIGPKMKLKNKTRILVTHSMYLPQVDVITVMSGGKI 846
Db 679 DITLPDPLSAVAHAGKHIFENVIGPKMKLKNKTRILVTHSMYLPQVDVITVMSGGKI 738
Qy 847 SEMGSYOEELARDGAFAEFLRTVASTEOBDAENGVTGSGPKGAQOMENGLVTDISA 906
Db 739 SEMGSYOEELARDGAFAEFLRTVASTEOBDAENGVTGSGPKGAQOMENGLVTDISA 798
Qy 907 GKOLQOLSSSSSYSDISRHNSHTAELOKAEAKEETWKLMEADKAQTOGVKLSVYDWY 966
Db 799 GKOLQOLSSSSSYSDISRHNSHTAELOKAEAKEETWKLMEADKAQTOGVKLSVYDWY 858
Qy 967 MKAIGLFIPLSLFLFPMCNHVSALASNYWLSLMTDDPIVNGTOEHTKRLSVGALGISQ 1026

```



```

Db      1196 LAVALGECVNCIVLFAALFAVISRSHSLAGLVGSLVSYSIQITAYLWLVMSSEMETNI 1255
Qy      1259 VAVRRLKEYSFETEEAAMQIOETAPSSWPOVGEVERNRNCLARRREDLPRLRIANTIN 1318
Db      1256 VAVRRLKEYSFETEEAAMQIOETAPSSWPOVGEVERNRNCLARRREDLPRLRIANTIN 1315
Qy      1319 GGEKGVIVGRTRGAKSSITLGLFRINSABEIIIDGINIAKIGLHDLRFKTIITIPDPV 1378
Db      1316 GGEKGVIVGRTRGAKSSITLGLFRINSABEIIIDGINIAKIGLHDLRFKTIITIPDPV 1375
Qy      1379 LFGSLRMNLDPFQSYDEEVMWLSLELAHLKDPVSALPDKLDHCAEGENTLVSQROLY 1438
Db      1376 LFGSLRMNLDPFQSYDEEVMWLSLELAHLKDPVSALPDKLDHCAEGENTLVSQROLY 1435
Qy      1439 CLARALLRKTILVLDATAVVDLETDLIQSTIRTOEDCTVTIARLNTIMDYRVI 1498
Db      1436 CLARALLRKTILVLDATAVVDLETDLIQSTIRTOEDCTVTIARLNTIMDYRVI 1495
Qy      1499 VLDKGEIOEYGAAPSDLLQORGLFYSMADAGLV 1531
Db      1496 VLDKGEYREGCAPSELQORGIIFYSMADAGLV 1528

RESULT 15
US-10-618-281-63
; Sequence 63, Application US/10618281
; Publication No. US20040219609A1
; GENERAL INFORMATION:
; APPLICANT: Day, Anthony G.
; APPLICANT: Beseff, David A.
; APPLICANT: Lyons, Eric H.
; APPLICANT: Yao, Jian
; TITLE OF INVENTION: Methods for Modulating Proteins Not
; TITLE OF INVENTION: Previously Known as Proteases
; FILE REFERENCE: GC773-2
; CURRENT APPLICATION NUMBER: US/10/618, 281
; PRIOR FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: US 60/395,325
; PRIOR FILING DATE: 2002-07-12
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 1303
; TYPE: PRN
; ORGANISM: Homo sapiens
; US-10-618-281-63

Query Match      61.3% Score 6068.5; DB 5; Length 1303;
Best Local Similarity 81.4%; Pred. No. 0;
Matches 1229; Conservative 17; Mismatches 57; Indels 207; Gaps 10;
Qy      22 WNTSNDFPTKCFONTVLVWVPCFLYMACPFYFLYLSRHRGYIOMTPLNKTITAGFL 81
Db      1 WNTSNDFPTKCFONTVLVWVPCFLYMACPFYFLYLSRHRGYIOMTPLNKTITAGFL 60
Qy      82 WYICMADLFSTFMRSGIGTILAPVFLVPTLLGTTLLATFLIQLERRKVOSSGIMLTF 141
Db      61 WYICMADLFSTFMRSGIGTILAPVFLVPTLLGTTLLATFLIQLERRKVOSSGIMLTF 120
Qy      142 WLVALCALALILSKIMTALKEADAVDLFRDITFYVYFSLLIQVLVSCSDSPLESET 201
Db      121 WLVALCALALILSKIMTALKEADAVDLFRDITFYVYFSLLIQVLVSCSDSPLESET 177
Qy      202 IHDNPNCPRESSASFLSHITFMWITGLIVRGYRQPLEGSDLSINKEDTSEQVVPVLVKN 261
Db      178 IHDNPNCPRESSASFLSHITFMWITGLIVRGYRQPLEGSDLSINKEDTSEQVVPVLVKN 237
Qy      262 KKECAKTRKQPVVYVSSKRPAPQKSSKVDANEVVALIVKSPQKMNSTLKVLYKT 321
Db      238 KKECAKTRN-----SSGSGSSCSANTALF-----PA--PTCHKSF 271
Qy      322 GPYFLMSFFPKAIDLMFMFGPOLIKLLIKFVNDTKAPDWQGYFYTVLLFVTACLQTLV 381

```

```

Db      272 QALSL-----LCRLILKFVNDTKAPDWQGYFYTVLLFVTACLQTLV 314
Qy      382 HOYFHICFVSQMKIKTAIVAVRKAIVITNSARKSTVGEIYNLMSVDAQRFMDLATYI 441
Db      315 HOYFHICFVSQMKIKTAIVAVRKAIVITNSARKSTVGEIYNLMSVDAQRFMDLATYI 374
Qy      442 NMMSAPLOVITLALYLMLNIGPSVLAVAVWMLMPVNAVMMKTTQYVAHMKSDNR 501
Db      375 NMMSAPLOVITLALYLMLNIGPSVLAVAVWMLMPVNAVMMKTTQYVAHMKSDNR 416
Qy      502 IKLMBELNGIKYLKUYAMELAFKDKVLAIROBELKVLKKSAYLSAVGTFTWCTPLVA 561
Db      417 IKLMBELNGIKYLKUYAMELAFKDKVLAIROBELKVLKKSAYLSAVGTFTWCTPLVA 476
Qy      562 LCTPAAVYVITDENNIIDACTAPVSLALFNILRPLNILEPVVSSIQVQSVSLRLRIFLS 621
Db      477 LCTPAAVYVITDENNIIDACTAPVSLALFNILRPLNILEPVVSSIQVQSVSLRLRIFLS 528
Qy      622 HEELEPDSIERRPVKKGGTNSITVRNATFTARSDEPTLNGITPSIPGALVAVVGQV 681
Db      529 GATSENGPWSRRKRG-----TROASFSVAEPVLCRFSITPSTIPGALVAVVGQV 581
Qy      682 CGKSSLSALLAEMDKVEGHVAIKGSVAAYVPOQAWIQNDSLRNIIFGCOLBEPYRSVI 741
Db      582 CGKSSLSALLAEMDKVEGHVAIKGSVAAYVPOQAWIQNDSLRNIIFGCOLBEPYRSVI 641
Qy      742 QACALLPDLLEILPSGRTIEIGKGVNLGGQKORVSLARAIVSNADITLFDPLSAVDH 801
Db      642 QACALLPDLLEILPSGRTIEIGKGVNLGGQKORVSLARAIVSNADITLFDPLSAVDH 701
Qy      802 VGHITFENVIGPQGMKNTRILIVTHSMSTLPQVDVITVMSGKISEMSYQELIARDGA 861
Db      702 VGHITFENVIGPQGMKNTRILIVTHSMSTLPQVDVITVMSGKISEMSYQELIARDGA 724
Qy      862 FAELRTVASTBQEQDAEENGTVGSGPKAKOMENGLVTDASAGKOLQRLSSSSYS 921
Db      725 -----S 725
Qy      922 GDISRHNSTALQKAEKKEETWKLMEADKQGVKLSVYDYNKAIGLFISFL 981
Db      726 CDL-----QVKLSVYDYNKAIGLFISFL 753
Qy      982 FMCNHYSAASNTWLSMTDDPIVNGTOERTKRLSVYGAIGISQIAYFGYSMAVSI 1041
Db      754 FMCNHYSAASNTWLSMTDDPIVNGTOERTKRLSVYGAIGISQIAYFGYSMAVSI 813
Qy      1042 ILASRCIAYDLIHSILRSPMSFPERPPSGNLVNRFSKELDYVSMIPEYIKMFGSLFNV 1101
Db      814 ILASRCIAYDLIHSILRSPMSFPERPPSGNLVNRFSKELDYVSMIPEYIKMFGSLFNV 873
Qy      1102 IGACIVILLATPIAIIIPPLGIYFPVORFYVASSROLKRLSVSRSPYSHNETLLG 1161
Db      874 IGACIVILLATPIAIIIPPLGIYFPVORFYVASSROLKRLSVSRSPYSHNETLLG 933
Qy      1162 VSVIRAFEEQERFIHOSDKVDENQKAYPSIYANWLAVERLECGVNCIVLPAALRAVIS 1221
Db      934 VSVIRAFEEQERFIHOSDKVDENQKAYPSIYANWLAVERLECGVNCIVLPAALRAVIS 993
Qy      1222 RHLISAGLVGLSVSYLQVTTYLNLVMSSEMETNIIVVERLKEYSFETEEAAMQIOET 1281
Db      994 RHLISAGLVGLSVSYLQVTTYLNLVMSSEMETNIIVVERLKEYSFETEEAAMQIOET 1053
Qy      1282 APPSSWPOVGRVFRNYCYRREDLDFVLRHINVTINGEKVIGVTRGAKSSITLGLF 1341
Db      1054 APPSSWPOVGRVFRNYCYRREDLDFVLRHINVTINGEKVIGVTRGAKSSITLGLF 1113
Qy      1342 RINESAEGEIIIDGINIAKIGLHDLRFKTIITIPDPVLSGSLRMNLDPFQSYDEEVM 1401
Db      1114 RINESAEGEIIIDGINIAKIGLHDLRFKTIITIPDPVLSGSLRMNLDPFQSYDEEVM 1173
Qy      1402 SLELAHLKDPVSALPDKLDHCAEGENTLVSQROLYCLARALLRKTILVLDATAVVD 1461

```

Db	1174	SLELAHLKDFVSALPDKLDHECAEGENLSVGORQVCLARALLRKTILVLDATAVD	1233
Qy	1462	LETFDILIOSTIRIQFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQORGLF	1521
Db	1234	LETFDILIOSTIRIQFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQORGLF	1293
Qy	1522	YSMAKDAGLV	1531
Db	1294	YSMAKDAGLV	1303

Search completed: December 15, 2005, 15:47:08
Job time : 171.832 secs

QY	1084	USMIPEVIMKMGSLFENVIGCIVTLTAPLAIIIIPGLIYEFVQSEFYAASSQGLRL	1143
Db	1132	DQIHPISTIECSRSTLLCVSALAVSIYTPFVLVPLALVAYCQIKQKRVASDLOOL	1191
QY	1144	ESVSRSPVYSHFNETLLGVSVLRAFEQORFIHQSLKVDENOKAYPYSIYANRWLAVRL	1203
Db	1192	DDTQQLPILSHPAETVEGJTTIRAFRYEARFOQKLEVTDSNNIASLPLTANRWLEVRM	1251
QY	1204	-----ECVGNCIYLPALPRAVIS--RHSLSAGIVGVSYSLSLOVTTYLLMWLVMSEEM	1255
Db	1252	ATPLPEQYETIACVVLILAAVTSISNSLRRELSAGLVGGLTYALAVSNYLLMMWVNLADM	1311
QY	1255	ETNIYAVERLKEYSEETERK-----APMOIQETAPPSMPQYGRVFPNVCILRYEDDDF	1308
Db	1312	ELQIGAVKRIHGLKLTAEASVEBGLAPSLI-----PKXWPDQKQIQNLSRYDSSLSLK	1366
QY	1309	VLRIHNTVINGEKQKVGIVRTGAGKSSLTLLGLFRINESAGEHIIIDGINIAKIGLHDLRF	1366
Db	1367	VLKHNVALISPGQKIGICGRTSGKSSPSLAFPRWDTFEGHIIIDGJIDIALPLHTRS	1426
QY	1369	KITTIPOBPVLFSGSLRMNLDPFQSGYSPBEWVTSLELHLKDFVSLDPDKLDHECAGEGE	1428
Db	1427	RLSIILQDPVLFSGTIRFNLDBPKCSOSTLEALEHILQKLVVRLPQGDALITTEGEE	1486
QY	1429	NLSVQGORQVLCIARALLKTKILVLDEATAVADLETDDLQSTIRTOFEDCTVLTIAHRL	1488
Db	1487	NFSQGORQFLCIARAFVARKTSIFIDENATASIDMATENILQCVMTAFADRTVITIAHRY	1546
QY	1489	NTIMDYTRYVILVDKEIIOEYAPSDLLQOR--GLPFSMAK	1526
Db	1547	HTIISADIVIVLKRGAILEPDKPEKLLRSKQSVESAFVR	1585

```

RESULT 4
US-11-090-439-26
; Sequence 26, Application US/11090439
; Publication No. US20050266442A1
; GENERAL INFORMATION:
; APPLICANT: Squillace, Rachel
; APPLICANT: Weiner, Michael P.
; TITLE OF INVENTION: Immortalized Human Tuberos Sclerosis Null
; TITLE OF INVENTION: Angiomyolipoma Cell and Method of Use Thereof
; FILE REFERENCE: 24318-502
; CURRENT APPLICATION NUMBER: US/11/090,439
; CURRENT FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: 60/556,344
; PRIOR FILING DATE: 2004-03-25
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 1581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-090-439-26

```

Query Match	21.2%;	Score 2100;	DB 7;	Length 1581;
Best Local Similarity	31.7%;	Pred. No. 4,8e-146;		
Matches 524;	Conservative 306;	Mismatches 567;	Indels 254;	Gaps 377;

```

QY      32 CFQNTVLVWVPCFYLMACPPFFL--YLSDHNGYQIMPTLNTKLTALGF---LIMVWC 86
      || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      26 CFVDALNVDPHFELLFTFFPLIFIGWSSQSKVHIHS-----TWLHPGNLWMLTF 79
      || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      87 ADLFPSFWERSGIGTLAPV-----FLVSPT---LLGITTLAFLIQLERKKVQSSG 136
      || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      80 MLFLVACEIAGEILSDGVTESHHLHLVMPAGMAFMAATVSV--YHNHIEFSNPKLLI 137
      || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      137 IMLTFTWLVALVALILSKINTALKEDAQVDLPFDITF-----YTFSLILLIQVLVS 189
      || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      138 ALLVVTTLAFL-----TKTIKFWALDLDAIGSQLRCLTGLGLVILVIGMTLLVE--VN 188
      || : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      190 CFSDSPLFSEF---IHDPNCPBESASTL-----SHITWMTTGLIVRGYRPLRGS 239
      || : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db	189	VIRRRRYEFFETPREVEXPEBDLOJGRFLQPFNLBSKGYWMMNFIKTAHKKPI---	245
Qy	240	DLWMLNKEDTSEQVUPVLK---NMKBECAKTRKQPVKVYSSKDPAQKESKVDANEE	296
Db	246	DLRAIGK-----LPIYMRALITYOJRLC-----	269
Qy	297	VEALIVSPQ--KENNPSLFKVLTYTFCGYPLMSFPFKAIHDMFMESGPOLKILIKFY--	353
Db	270	FDQVORXDIQGTQGARAIWQALSHAFGRRLVLSSTFRILADLQFAGLCTFGIVDHLGK	329
Qy	354	-NDTKAPDMQ-----GYFYVTLTPVTAQLOTLVHQFPIFCVSGMRIT	397
Db	330	ENDVFOPEQTQGLGYFVSSQGFELNAYVNLTLFLALLQFTFLQFQASYYVAIFEGINRG	389
Qy	398	AVIGAVYRKA--LVIITNSARKSFTVEBIVNIMSVDAORFMDLATYINNIMWAPLOVIAL	455
Db	390	AIQTKIYKIMHLSSTNLSMGEMTAGQCNLVALIDTQMLMFFELCPMLAMPQOIIYGV	449
Qy	456	YILMLNLGSPVLGAVVWVLPVNPVNAWAMTKTYQVAHMSKNRILKLMNEILNGIIVL	515
Db	450	ILLYYIIGVSLIGAAVILILAPQYFATLQSGAOSTTSEYNSERLQTNEMTRGIKLL	509
Qy	516	KLYAMELAFKQKYLAIROEBLKVLKKSAYLSAVSTFWVCPPFLVALCTFAVYLTIDENN	575
Db	510	KLYAMENIFRRVETRKRKMTSLRARIYISISIFMNTALPILAAVLITFGVHSFFIEA	569
Qy	576	ILDQATAFVSLAFNIIIFPPLNIIIPWYISSIVQASVSLKRLRIFLSHEELEDST----	630
Db	570	DFSPSVAFASLSLPHILVTPLFLSSVYRSTVKALVSQKLSFEFLSAEIRBEOQAPHEP	629
Qy	631	-----ERRPVK-----DGGGTN--SLIVRNATFPWA	654
Db	630	TPQGBASKYQAVPLRVNKRKRAPBEDCGLTGPLQSLVPSADGADNCVOIMGGYFTWT	689
Qy	655	RSDDPTLNGIFESFPEGALVAVWQOVGGCKSSLALSALAEADYEG-----	700
Db	650	PDGIFTLSSNITIRIPRGQITWIVGQVGGCKSSLALLAGEQKYSGAVFMSLPSDEIGE	749
Qy	701	-----HVAIKGSVAVYVPQAOIWNDSIRENIIIFGCODEEPPYRSVIOACALLP	748
Db	750	DSPERERATDLDLRKRGPVAVYASQKPLNATAYENIIFESPKNQKQKMWITEACSLQP	809
Qy	749	DLEIIPSGDRFEIGEGVNSLGGQKORVSLARAVYSNADITLFPDPLSAVDAAHGXHIFE	808
Db	810	DIDILPHEDQOQIGERGINLSSGGQRORISVARALYOHANVVELDDPFALDIHLSDMIQ	869
Qy	809	NVIGRKGMLK-----NKTRIIVTSMSTLPOVDYIYVMSGKISBMGSYOBLARDGAFA	863
Db	870	-----AGLEILLRDDKRTVVLVTHKQVLPADMIITAMKDGITQREGLTKQFQMSBEOJF	924
Qy	864	EFLRTYASTEOBODAEENGVTGVSQPGKEAKQEMENGLVYTSAGKOLQROLSSSSYSGD	923
Db	925	EHWKTIWN--RODQBELKETV-----BKKAITEPQG-----LSRAMSRRGL---	965
Qy	924	ISRRHNSTAELQKAEKKEETWKLMEADK-----QTGQVLSYVWYMKAIGLFISLS	978
Db	966	-----IQDEEBEEBAESEDNDNLSMLHQRAEIPWRACAKYLSAGIILLSLTL	1015
Qy	979	IFLEPMCHVNSALASNYMLSLMTDDPIV-----NGQOERT-----KXLSYGAALGI	1024
Db	1016	VFSQLLKTMVLAIDWYLMKWTDSALTTPAARNCSLSQECTLDQTYVAMFYVLCISGI	1075
Qy	1025	SOGIAVFGYSMAVSGIGILASRCUHVLDLHSLRSPMSFPFERTSGNLVNRFSKELDTVD	1088
Db	1076	---VLCIVTSYVTEWTEGLKAKAKRLHRSLLNNIILAPMRFFETPLGSIANFSSCMTID	1133
Qy	1085	SMIPEVIMFGSLFNVIACIVILLATPIAAIIIPGLIYIFVQOFRYVASRQKLE	1144
Db	1133	QHIBPTECLSRSTLLCVSALAVISYVPEVLALLPLAVVCYFOKRYVASRDLQOLD	1192
Qy	1145	SVSRSPVYSHRENTLGVSVIRABEQGRFHQSDKVDENQAKAYBSIVANRMLAVYLE	1204
Db	1193	DTQOPLLSHAETVEGLTIRAPRYERFOOKLEIYDSNNIISLFLTANRMLTEVME	1253

Qy	1205	CWGNATIVFAALFAVTS--RHSLSGLVGLSVSYSLQVTTYLTANMLVRMSSEKNETIVAVE	1265
Db	1253	YIGACVULIAAVTSTISNSLHRLSLGVLGCLTYLTAAMSNTLMMVNRNLADIELQIAVK	1312
Qy	1263	RLKEYSETEKE-----APWQIQETAPSSWPQVGRVFRNYCRLYREDDLPVLRHINVT	1316
Db	1313	RHGLLKTTEAESYEGELALPSSLI-----PKMWDQCKIQIQNLSVARYDSSLKRVLGHVVAL	1367
Qy	1317	INGGKNGVIVERTGKSGSLTGLFRINBSAGEIITIDGINAKIGLHDKRKTIITFOD	1376
Db	1368	ISPQKICIGCGTGGKSSFSFLAFPRMVDTFEGHIIIDIGIDIRKLPHTLPSRLSIIOD	1422
Qy	1377	PVLFGSGLRMMNDPFSQXDEEVMWTSLELAHLKDFVSLAPDGLDHECAGEGENLSVGORO	1436
Db	1428	PVLBSGTRFPLNDPERKCSDSLTMALQIAKLKYKALPGGLDAIITTEGENSSQGORO	1487
Qy	1437	LVCLARALLRKTKILVLDENAVDAVDLETDLLIOSTIRQFEDCYVLTIAHRNLTIMDYTR	1496
Db	1488	LFCLARAFVTRKTSIFIMDEATFASIDMATENITQKVMTAFADRTVVTIAHRVHTLSADL	1547
Qy	1497	VIVLDKGEIOEYGAPSDLLQOR-GLFYSMAX 1526	
Db	1548	VIVLRGALILEFDEKPEKLLSRDVSFAFSVR 1578	
RESULT 5			
US-10-995-561-526			
; Sequence 526, Application US/10995561			
; Publication No. US20050272054A1			
; GENERAL INFORMATION:			
; APPLICANT: CARGILL, Michele et al.			
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH			
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF			
; TITLE OF INVENTION: DETECTION AND USES THEREOF			
; FILE REFERENCE: C1001559			
; CURRENT APPLICATION NUMBER: US/10/995,561			
; CURRENT FILING DATE: 2004-11-24			
; NUMBER OF SEQ ID NOS: 85702			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 526			
; LENGTH: 1404			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-995-561-526			
Query Match 13.5%; Score 1341.5; DB 6; Length 1404;			
Best Local Similarity 27.8%; Pred. No. 2.2e-90;			
Matches 369; Conservatve 242; Mismatches 473; Indels 243; Gaps 34;			
Qy	32	CFQNTVLVWVPCFYLMACFPFYFL-YLSRHDGVIQMTPLNKTATLGF-----LIMTICW 86	
Db	26	CFVDALNVVPVFLFITFTPLFTIGWGSQSSKVHIHS-----TWLHPGHNLRWILTF 79	
Qy	87	ADLFYSFEMERSRGIFLAPV-----FLVSPF-----LIGITLLATFLIQLERRKGVQSSG 136	
Db	80	MLLFVLVCEIAEGILSDVETSHHHLMLVMPAGMAFMAVTSV--YHNITISNPKLLI 137	
Qy	137	IMLTFVLVALYCALALRSKMTALKEDAOVDLPDITF-----VYFSLLLIQLVLS 189	
Db	138	ALLVYMTLAPF-----TKTIKVFKFLDHAIGFSQILFCUTGLVILVILYGMILLVE--VN 188	
Qy	190	CFSDRSPLFSETIHDNPNCPS-----SASFSLRTLPFWMITGLIYGRQPLEG 238	
Db	189	VIRRRRYIFFTPREVKE-PEDDLODLGVRFQPFVNLSSKGTYYMMNAFITAAKRP-- 245	
Qy	239	SDLSLANKEDTSBOVVPVLVK--NWKECAKTRQVPKVVVYSKDPAPQYESSKVDANE 295	
Db	246	-DLRAIGK-----LPIAMRALNTVYORLC-----E 268	
Qy	296	EVEALLIVSPQ-XEENPSLRFVLVYKTRGPPYFLMSFFFAIHDLMMFSGPOLIKLLIKV- 353	
Db	269	AFDQVQRKDIOGTQGARAIWQILSHAFERRVLVLSFTFIIADLDFGAPRLICFGIVHDLG 328	

```

QY 354 -NTRKADMC-----GFTYVLLFVTAQCTDTVLVHQFHI CFVSGMIX 396
Db 329 KENDFPQKTOFLGYFVVSQOEPFLANAVTLAFLALLORTFLQASVYAIETGINIK 388
QY 397 TAVIGAVVRKA--LVITNSARKSSTGVEI VNLMSVDAORFMDLATYIMIMVASPQVILA 454
Db 389 GAIQTKIKMLHISTLSMLSGEMTAAQICNLVVIDTQMLMFFLPCNLWAMPQOIIYG 448
QY 445 LYLMLNLGPEVLGAVAVMLVVPYNAVMAKTKTYQVAMHKSJONRIKLANEIIINGIKV 514
Db 449 VILLYYIIGVSALIGAAYIILLAPQYVVAATKLSQAQRSTLEYSNERIKOTEMIRGIKL 508
QY 515 LKTVAMELAFQDKYLAIRQOELKYVLKKSAYLSAVGTFFWVCPPLVALCTPAVVYITDEN 574
Db 509 LKTVAMENIFETRETTTRKEMTSLRAPAYITSIFENNTAIPLAAVLTIVGHVSFFKE 568
QY 575 NILDQAFAVSLAENILRPLNLLPMVSISSIVQASVLKRIARFLSHHELEPDSI----- 630
Db 569 ADFSPSVAFAHLSLPHILVTFPLILSSVRSYTVALVSVQKLSBELSAAELREBCAPHB 628
QY 631 -----ERRPVK-----DGGCIN-SITVRNATTW 653
Db 629 PTPQGPASKYOAVPLRVNRRKPARREDRGTLGTLQSLVPSADGADNCVQIMGVEFTW 688
QY 654 ARSDPPTINGTFTSIPEGALVAVVGVGCGKSSILSLAAMDKEG-----700
Db 689 TPDGIPTLSNITIRIPRQGLTMIVGQVCGGSSLLAALGEMOKSGAVFMSSLPDSEIG 748
QY 701 -----HVAIKGSVAVYVPOQAMIONDSIRENILEGCOLLEBYRVSVOACALL 747
Db 749 EDPSPERTATLDRIKRGPIVAVASQKRWMLNATYEENIIFESPNNKRYKMYIEACSLQ 808
QY 748 PDLBILPSGDFTEIGEKVNLSSGGQKQRFVSLARAVYSNADYIFLDDPIISAVDAYGKIF 807
Db 809 PDIDILPHGDOQTOIGERGINLSSGQQRQISVARALYCHANVFLDDPSALDIHLSHLM 868
QY 808 ENVIGPGRGMK-----NKTRLVTHMSVYLQOVVITYMSGKISMSQYOBELLARDQAF 862
Db 869 Q-----AKILELADDDKRTIVLVTHKQYLFHADWTIMKQGTIOREBGLTDFQPSRBCOL 923
QY 863 AEFRLTYASTEOBDAENGVTGVSGPGEKAKOMENGMVLVDSAGKOLOROLSSSSSYSG 922
Db 924 FEHMKITLMN-RQODELEKETV-----ERKATEPPQG-----LSRAMSNDGL-- 965
QY 923 DISRHNHSTAELOQAQAKKETWKLMEADKA-----QTGVYKLSVYMDYKATIGLFISFL 977
Db 966 -----LQOEBEEBEEBAESEBDDNLSMLHQAEALPMPACAKYILSAGILLLSL 1014
QY 978 SIFLMCNHVASLNSWYLSMTDDPIV-----NGTOEHT-----KYRLSYGALG 1022
Db 1015 LVFSQOLKHMVLVAIDYMLAKMTSALTITPPAARNCISLSQECTIDDQYIVAVVFLVSLG 1074
QY 1024 ISQGIAGVGYMAVSIIGIILASRCILAHVLDLSILRSPMSFEERTPSGNLVNRFSEKLDTV 1083
Db 1075 I---VLCIVNYSVYEMWGLKVAKRHLHSILNRITLAPMRPFETTPPLSILNRFSDCTI 1133
QY 1084 DSMIPEVIKMFMSGLFNVIACIYIILATPIAAITIPPLGIYTFVQORFYAASBQKRL 1143
Db 1132 DQHLPSLTLECSRSTLLCVSALAVISYTPVFLVALPLAIVCYCIFIQCYFVASRDLQOL 1191
QY 1144 ESVSRPVSYPHFNTLIGVSVYIRAFEEBOERITHOSDKVDENQCAVYPSIYANMIAVRL 1203
Db 1192 DDTTQOLPLSHFATVEGLTITIRAFYBARQOKLEBTDSNNIASLFLTAANRMLBVRM 1251
QY 1204 ECVGNCTI 1210
Db 1252 -AIRCM 1257

```

```
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C0001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 969
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-969

Query Match      9.8%; Score 969; DB 6; Length 410;
Best Local Similarity 50.8%; Pred. No. 8e-64;
Matches 197; Conservative 66; Mismatches 121; Indels 4; Gaps 2;

QY 495 MRSKDNRIKLNELINGIKVLKLVAMELAFKDKVLAIRQSEELKVLKKSAYLSAVGTFTWV 554
DB 1 MRQDSRAKLTSSILRNKSKTKFHMEGAFIDRVLGIRGELGALRTSGLLPVSLSVRFQ 60
QY 555 CTPELVALCTPAAVYVTTIDENNILDAQTAFVSLAFNLIPLNLIPLNVISSIVQASVSLK 614
DB 61 VSTFLVAVFVAHTLVAE-NANNAEKAFVTLVNLINKAQAFLPFSIHSLVQARVGF 119
QY 615 RLRLFLSHLEELPDSIERPKYDGGTMSITVRANKTFWARSDDPTLNGITFSPLEGALV 674
DB 120 RLVTFLCLIEVDPGVAVSSSGSAAGKDCITIQSTFAMSSQSPCLHRIMLVPOGCLL 179
QY 675 AVVGQVGGKSSLSALIAEMDKVEGVAIKGVAVVYVPOQAMIONDSRENILLEGCOLLE 734
DB 180 AVVGPVGGKSSLSALIGELSKVEGFVISGAVALYQZAMVQTTSVENVCQGEIDP 239
QY 735 PYRVSVIQACALLPDLLEILPSGDRTELEGKVNLSGGQKQVSLARAVYSNADITLPDP 794
DB 240 PWLERVLEACALQPDVDSFPFGIHITSIGEGMNLSSGQKQLSLARAVYRRAAYLLDDP 299
QY 795 LSAVDAYGKATFENVIGPKMKLNKKTITLVTHSKSYLPQVDVITVWGSGKISEMGSTOE 854
DB 300 LAALDAHVGQHVFNQVIGPGLLOCTRTITLVTHALHILPOADWITVLNGAIAEWGSYOE 359
QY 855 LLARDGAFAEFLRYASTEOBDAENG 882
DB 360 LLQKKGALVCL--DQARQPDGRGEGG 384

RESULT 7
US-10-613-744-14
; Sequence 14, Application US/10613744
; Publication No. US20050272093A1
; GENERAL INFORMATION:
; APPLICANT: MacKinnon, Roderick
; TITLE OF INVENTION: The Rockefeller University
; TITLE OF INVENTION: Assays for Screening Compounds Which Interact With
; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
; TITLE OF INVENTION: Channel Proteins, and Uses Thereof
; FILE REFERENCE: 018512-002901US
; CURRENT APPLICATION NUMBER: US/10/613,744
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US/09/275,252
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: US 09/045,529
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/054,347
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: WO PCT/US99/06307
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 14
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-613-744-14

Query Match      8.9%; Score 883.5; DB 6; Length 407;
Best Local Similarity 48.7%; Pred. No. 1.5e-57;
Matches 167; Conservative 74; Mismatches 95; Indels 7; Gaps 4;

QY 1551 LTRLAEDPAEPRY----RTERRARFYSKGCNCAVAKNIRQGEQFIDVPTTIVDLKWP 1606
DB 16 LPRQARDL--PRHISRDTRKRIQRYVKKQCNHNGVREYRITLITFTLVLDLKW 74
QY 1607 HTLLIFTMSPFLSCWLLPAMVWMLIAFAHGDLPAGEGTN-VPCTVSIHSFSSAFPSIEVQ 1665
DB 75 FNLIIFVMVTVTWLFPFGMTIMWLAIYRGMDHDESDPWPPTCVNNLNGVSAFLFSIETE 134
QY 1666 VTIFGGRMWTBECPLAILIIVONIVGLMINAIMLCCIFPKTAQAHRAETLIFSRAV 1725
DB 135 TTIGYGRVITDKCPREGIILLILOSIVGSIYNAFVQCMFVKISQPKRAETLVFSTHAV 194
QY 1726 ITLHGRGLCEMLRYGDLRKSMITISATIMQVYRRTTSPREGVVPILHVDIPMENVGGNG 1785
DB 195 ISMDGKLCIMFRVGDLRNSHIVASIRAKLIKSKQTSSEGFPLNQTIDINVGYYTGDDR 254
QY 1786 IFVLAPLIYHVIDSNSPLYDLAPSDLHHODEIIVILEGVETGTTQARTSYLADE 1845
DB 255 LFLVSPILISHIEINQSPFWEISKQOL-PKEELEIIVILBGMVATGMTQCARSSYITSE 313
QY 1846 ILWGQRFVPIVAEEDGRYSVDYSKFGNTIKVPIPLCTARQOLDE 1888
DB 314 ILMGYRFPTVLTLEDGYEVDYNSFHEHYETSTSLAKELAE 356

RESULT 8
US-10-995-561-967
; Sequence 967, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C0001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 967
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-967

Query Match      8.1%; Score 799; DB 6; Length 268;
Best Local Similarity 56.3%; Pred. No. 1.3e-51;
Matches 151; Conservative 56; Mismatches 61; Indels 0; Gaps 0;

QY 1264 LKYESETEKAPWQIQETAPSSWPQYGRVEFRNYCIARYEDDLFVLRHINVTNGEKV 1323
DB 1 MQDYAWTPKEAPWRLPICAQPPMPQGGQIEFRDFGRLRYRDELPLAVQVSFKIHAGKV 60
QY 1324 GIVRTAGKSSLILGPRINESAGEIITDGINIAKIGLHDLAFKTIITIPQDVLPSGS 1383
DB 61 GIVRTAGKSSLASGLRLQEALEGGWIDGVPIAHGHTLRSRSLIIPQDVLPSGS 120
QY 1384 LRNMIDPFSSQYSDDEVVTSLELAHKDFVSAIPLKDLHECAEGGENTSVGQRLVCLARA 1443
DB 121 LRNMIDLLQHSDDAIAALETTVQLKALVASLPQLOLYKCKADREDDSVGQKQLCLARA 180
QY 1444 LLRTKILIVDEATAVADLETDLLIQSTIRTOPEDCVLTITAHRLNTIMDYTRVIVLDKG 1503
```

Dh 181 LARKTOILLDEATAAVDPGTELOQMALGSMFAQCTVLLIYHRLRSVMDCARVLVMDKG 240
Qy 1504 EIOEYAPSDLLQORGLFYMAXAGIV 1531
Dh 241 QVABSGSPAQLLAQKGLFYRLAQSGLV 268

RESULT 9
US-10-613-744-13

; Sequence 13, Application US/10613744
; Publication No. US20050272093A1
; GENERAL INFORMATION:
; APPLICANT: MacKinnon, Roderick
; TITLE OF INVENTION: The Rockefeller University
; TITLE OF INVENTION: Assays for Screening Compounds Which Interact With
; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
; FILE REFERENCE: 018512-002901US
; CURRENT APPLICATION NUMBER: US/10/613,744
; PRIOR FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US/09/275,252
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: US 09/045,529
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/054,347
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: WO PCT/US99/06307
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-613-744-13

Query Match 7.9%; Score 786.5; DB 6; Length 391;
Best Local Similarity 46.2%; Pred. No. 1.9e-50;

Matches 150; Conservative 75; Mismatches 89; Indels 11; Gaps 6;

Qy 1564 RTERRARFVSKKNCNVAHNIREQRF--QDVFTLVLDKMPHTLLFTMSFLGSL 1621
Dh 35 RSRQ-RARLVSKGRCNIEFGVDAQSRFIFVDIMTVLDLKKRYKMTVITATFLSWMF 93
Qy 1622 LFNAMWMLIAPAHDL-----APGEGTAVPCYTSIHSSSAPLSEIYQVITGFGGRWTE 1677
Dh 94 LFGLLWYVAVVHNDLPDEFYPPDNR--PCVENINGMTSAFLFSLFQVITGYGFRFVTE 151
Qy 1678 ECPALAILLYQNTVIGLMINAMIGCTFMKTAQARRAETLIFSKHAVITLHGRLEPML 1737
Dh 152 QCARATIFULLFQSLIGYIINSFMCALILAKISRKPKAKTITTSKNAVISRGGKLCILI 211
Qy 1738 RVGDLRKSMTISATIHQVVRKTSPPGEVVPPLHQVDI PMENGVGNGCIFLVAELIYHV 1797
Dh 212 RVANLFRKSLIGSHIYGLTKLTITTPSGETIILDQTNINPVVDAGNENLPIISPLIYHI 271
Qy 1798 IDSNSPLVDLAPSLDHLHODLEIIVIEGCVETTGITTOARTSTYLADEIILMGQFPPIVA 1857
Dh 272 IDHNSPFHMAAETL--SQODFELVFLDGTVESTATCOVTSYVPEEVLMGYRFVIVS 330
Qy 1858 E-EDGRYSVDYSKFGNTIKVPTPLC 1881
Dh 331 KTKGKTRVDFPHNGKTVAVETPHC 355

RESULT 10
US-10-793-626-326

; Sequence 326, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US

; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 326
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-326

Query Match 5.2%; Score 514.5; DB 6; Length 548;
Best Local Similarity 27.1%; Pred. No. 3.2e-30;
Matches 157; Conservative 110; Mismatches 211; Indels 101; Gaps 16;

Qy 1002 DRYNGTQERTKRLSYGALGISQGIYAVGYMAVSIIGILASRCLHVDLHSLTASPM 1061
Dh 12 DGVIN--NHSLTNQEKFSHGVAGIALF-----IFLIVRPPI 47
Qy 1062 SFEE---RTPSGNLVNRFSEKL-----DTYDSMPEYIK-----MFMG 1096
Dh 48 EFRQYLAQMTNSKILIDIRKQNLHQLSAYRFYANNOYGQVSRVINDEYQTKDPILT 107
Qy 1097 SLENVIGACIVILLAPI-----AAIIPPLGL--IYFVQGFYVASSRQLRLE 1144
Dh 108 GLNMIMDCITIIITIALSIMFPLDVKLFPALIFFPYILIVYFFPGL-----RKLTVRV 162
Qy 1145 SVSRSPYSHFNFTLGVSVITRAFEBOERTIHQSDLVKNDENQKARYYSIYANRLAARLE 1204
Dh 163 SQALAEVQGFLEHVGQMSVYKSFALIDENAKNFD--NNKQFLQRAFOHTRWNAYSFA 219
Qy 1205 CV-----GNCIYLFALFVVISRHSLSAGLVGSYSLOVTTYLMVLRMSSEMETNI 1258
Dh 220 AINTVTDLGPILYIVGVSYLAIT--GSTTVGTTLAAFGYLEQDLPRLRLVSSFTTLTQSF 278
Qy 1259 VAVERLKEYSBTEKE-----APWQIQETAPSSWPQGVREFRNYCLRYRBDLPVL 1311
Dh 279 ASMDRVQLMDEDDIDKNGIQAQPIKSK-----QIDDKHVSFYNNKEKVELH 328
Qy 1312 HINVTINGEKRGVIGVRTGKSSLTIGLFRINESABGELIITDGINAKGLMDLRKIT 1371
Dh 329 DINTLTINKGRTVAFVGGSGGKSTLINLRFYDVYQGEILIDHNKVDPLTGLSRNQIG 388
Qy 1372 IIPQDPVLFSGSLRMLN---DPSQYSDSEVWTSLELAHLKDFVSALPDKLDEHCAEGE 1428
Dh 389 LVQQDNILFSDTYKENILLGRP--DATDDEVBAKMAAHADFTSLNPNGYDTEVGEGV 446
Qy 1429 NLVGQROVLCLARALLRKTILVLDATAVADLETDLLIOSTIRTOFEDCTVLLIARL 1488
Dh 447 KLSGGQGORSTARIFLNNPVLILDEATSALEDSALIIQELMDVLSKORTLLIVARL 506
Qy 1489 NTIMDTRVVVLNKGSIQEGAPSDLIQQRGLF---YSM 1524
Dh 507 SITTHADRIYVMEGRIVETGTHQOLINKGAAVEHLYSI 545

RESULT 11
US-10-467-657-1346

; Sequence 1346, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASTIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657

;; CURRENT FILING DATE: 2003-08-11
;; PRIOR APPLICATION NUMBER: GB-0103424.8
;; PRIOR FILING DATE: 2001-02-12
;; NUMBER OF SEQ ID NOS: 9218
;; SOFTWARE: SeqwIn99, version 1.04
;; SEQ ID NO 1346
;; LENGTH: 622
;; TYPE: PR1
;; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1346

Query Match 5.0%; Score 495.5; DB 6; Length 622;
Best Local Similarity 28.7%; Pred. No. 9.6e-29;
Matches 154; Conservative 106; Mismatches 213; Indels 63; Gaps 13;

QY 1022 LGISGIAVFGYSMAVISGIGLASRCLHVDLHSLRSPMSFPERPSGNLVNRFSEKLD 1081
D 108 LVTARICRFTSTYLTMTWVSVMITSKIKDMFAKKLTLSSKHQETPSGTVLMMNLTE 167
QY 1082 TVDSMIEVIMFMGSLFNVIGACIVLL---ATPIAIIIPPLGIYFVQRFYVASS 1137
D 168 QSVNASADIFVTLRDTWITGTLFTVLLYLMWQSLIVLWMPLLSL---SRYY---R 220
QY 1138 RQKRLSRSVSRSPVYSHN---ETLLGSVTRAPEQGRFTHQSDL-----KYDE 1184
D 221 DRLKHVLSDSQKSGIGTMNVIAETHQGRVVKLFNGQAQANRPDAVNRITVRLSKITQ 280
QY 1185 NQKAYP--SIVARNMLAVRLECVGNCIVLPALPAVSRHSLSAGLVGSYSLOVTT 1242
D 281 ATAAHSPSELSIASTALAV-----VIFALMQSGNGTTTIGEPMAFIVA-MLOWYA 330
QY 1243 YLNLVNRSSSEMETNIVAVERLKEYSETEKE-----APWQIOETAPSSWPQVGRVER 1296
D 331 PIKSLANISIMQMTMFLAAGVCAFLDTPPDQDKGTALPQRYE-----GRISFR 379
QY 1297 NYCARYEDDLPVLRHINVTINGEKVGIYGRTGAGKSLTLGLFRINESAGEIITDGI 1336
D 380 NVDEYRSDDGIRKALDNFNLDIRQGERVALVRSRSGSKSTVNNLPRFEPAGNCTIDGI 439
QY 1357 NIAKIGLHDLRFKTIITIPQDPVLPFSGSLRMNLDPSQYS-----DEEVTSLELHAKDF 1411
D 440 DIATIKIDCLRAQALVSGQDFLEPDTTFENV---KRSRPDAGEAVLSALQANLUSL 495
QY 1412 VSALPDKLDHCEAGGENTLAVGQRLVCLARALLRKTILVDEATAVADLETDLIGST 1471
D 496 IDASPLGLHQPRTIGSGNSLSCGQGRVAVARAILKADAILLDEXTSALDNESERLVOQA 555
QY 1472 IRTQPEDCTVTIARLNTIMDYTRVYLDKGEIOEYGAPSDILQORGLFTSMAD 1527
D 556 LERLMENRTGIIVAHRLTTVESADRIIVMDGKIIEOGTHDQLMFONG-YVTMLRN 610

RESULT 12
US-11-082-389-148
; Sequence 148, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-131CPN
; CURRENT APPLICATION NUMBER: US/11/082.389
; CURRENT FILING DATE: 2005-03-16
; PRIOR APPLICATION NUMBER: US 09/603024
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/143262

;; PRIOR FILING DATE: 1999-07-09
;; PRIOR APPLICATION NUMBER: US 60/151281
;; PRIOR FILING DATE: 1999-08-27
;; PRIOR APPLICATION NUMBER: DE 19930487.4
;; PRIOR FILING DATE: 1999-07-01
;; PRIOR APPLICATION NUMBER: DE 19930489.0
;; PRIOR FILING DATE: 1999-07-01
;; PRIOR APPLICATION NUMBER: DE 19931549.3
;; PRIOR FILING DATE: 1999-07-08
;; PRIOR APPLICATION NUMBER: DE 19931550.7
;; PRIOR FILING DATE: 1999-07-08
;; PRIOR APPLICATION NUMBER: DE 19932134.5
;; PRIOR FILING DATE: 1999-07-09
;; PRIOR APPLICATION NUMBER: DE 19941379.7
;; PRIOR FILING DATE: 1999-08-31
;; Remaining Prior Application data removed - See file wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 446
;; SEQ ID NO 148
;; LENGTH: 549
;; TYPE: PR1
;; ORGANISM: Corynebacterium glutamicum
US-11-082-389-148

Query Match 4.7%; Score 468.5; DB 7; Length 549;
Best Local Similarity 26.9%; Pred. No. 7.7e-27;
Matches 141; Conservative 110; Mismatches 201; Indels 73; Gaps 18;

QY 1031 FGYSMAVISGIGLASRCLHVDLHSLRSPMSFPER-----TPSGNLVNRFSKELTVD 1084
D 50 FGSKLSRVRG-----DLRSALFGKVVNFSEREMQGFAGS--LITRINDVQOVQ 98
QY 1085 SMIEVIMFMGSLFNVIGACI-----VILATPIAIIIPPLGIYFVQRFY 1133
D 99 MLVQMTSTLMSAVALIGGIIAMVRQDLGSLWLVVSIPLIIV---ALIIYRAMPPLF 155
QY 1134 VASRQKRLSESRSRSPVYSHFNETLIGSVYRAF---BEQEFTHQS-----DLKVD 1183
D 156 QTMOKRIDRINQIIR-----EQLTGIRITRAFVEDERERFTTASXQVADIGVRTG 207
QY 1184 ENQAYYPSIYARNMLAVRLECVGNCIVLPALPAVSRHSLSAGLVGSYSLOVTTY 1243
D 208 NMLAMFPAVA---LIMNISAVA---VIFGAFQVSGSET-QIGTLPALFQIYMQLIMG 259
QY 1244 LNLVNRSSSEMETNIVAVERLKEYSETEKEAPWQIOET-APPSSWPQVGRFERNYCLRY 1302
D 260 VMMAAFVFWVPRAAVAGADRIGEVLETPGV--QAPTPAQPT--SAGEIVFNNATPAY 315
QY 1303 REDIDFVLRHINVTINGEKVGIYGRTGAGKSLTLGLF-RINESAGEIITDGINAKI 1361
D 316 PGADDPVLNNVSPFVAGSTTALIGSTGSGKTTL-IGLVRLPFTVTEGDVVDVDTVREF 374
QY 1362 GLHDLRFKTIITIPQDPVLPFSGSLRMNLDPSQYSE-----EVTSLLELHAKDFVSLAP 1416
D 375 EPLKLMDRIGLVPQKSLFSGTASNL---RIGNEBATETQWQALAIQAADPVBMBP 430
QY 1417 DKLDECAEGGENTLAVGQRLVCLARALLRKTILVDEATAVADLETDLIGSTIRTOF 1476
D 431 EGLDSEIAGGTVNSGQGRQRLAIRALLKQPEIYIYIDDSFSAIDVSTDAALRALSTNL 490
QY 1477 EDCVTLTARLNTIMDYTRVYLDKGEIOEYGAPSDILQORGLF 1521
D 491 PDATKLIIVQRVSTIRBADQIVVDNGEVVGIGTHMLNTMTCGY 535

RESULT 13
US-10-995-561-575
; Sequence 575, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF

Db 247 LSLFLVLVLSLGEMLPFFTRLTWILLQDGSADFTFTRNLTLMSILTLIASAVLEF--- 303
QY 1035 MAVSIG-GILASRCIHY-----DLHSHILRSFMSFFERTSPSGNLVNRFSKELDTVDSMI 1087
Db 304 ----VGDGITYNNTWGHVHSHLQGEVFGAVLRQETFEFQONOTGNIMSKRVETDSTLSDSL 359
QY 1088 PEVIKMFMSLFENVIGACIV-ILF-----ATPIAIIIPPLGLYFFVQRFYVASSRQLK 1141
Db 360 SENLSTLFIMYL--VRGLCLGIMLMGVSFLTMVTLITLPLFLPKVKGMVQLLEVCVR 417
QY 1142 RLESVSRSPVSHNETLLGSVIRAFEBQRFTHQSDLKVE---NOK--AYPSIV 1194
Db 418 --ESLAKSSOYA--IEALSAMPVTRSEFNEGEAQKFERKLOEIKTLNQEAVAY---A 469
QY 1195 ANRW-----LAVRLECVNCCIYLFALFAVISRHSLSAGLVGLSVYSILOVTTYLNM 1246
Db 470 VNSWTTGISGMLLVGILYIGQGLVTSGA-----VSSGNLVTFVLVYOMQFTQAVEV 520
QY 1247 LVRRSSEMETNIVAVERLKEYSETEKEAPWQIQETAPPS---WPQYGRVEFRNYCLRY 1302
Db 521 LLSIYPRVQKAVGSSEKIFEYLDRTPR-----CPSGLITPLHLEGLVQFDVSPAY 572
QY 1303 --REDLDPVLEHINVTINGEKVGVGERTGAKSSLTGLFRINESAGEIITIDGINIAK 1360
Db 573 PNRPDV-LVLOGLFTFLRPGEVTAIVGPNSGSKSTVALLQNLVQPTGGQLLDGKPLPQ 631
QY 1361 IGLHDLRFKITIIPQDPVLPFSGSLRMNL--DPFSQYSDDEVWTSLELAHLKDPVSALPDK 1418
Db 632 YEHRYLHRQVAAVGOEPVFGRSLOENIAYGLTQKPTMEBITAAAVKSGAHSFISGLPQG 691
QY 1419 LDHECAGEGENTSVGQRLVCLARALLRKTILVIDENTAAMD---LETDDLIOSTIRT 1474
Db 692 YDTEVEDEAGSQLSGQROAVALARALLRKCVCVLIIDATSAIDANSQLOVEQLYES--P 749
QY 1475 QFEDCTVLTIAHRLNTIMDYTRVYVLDKGEIQEYGAPSDLLQQRGLFYSMAK 1526
Db 750 ERYRSRVLLITQHLSLVEQADHILFLEGGAIREGGTHQOLMEKKGCVWAMVQ 801

Search completed: December 15, 2005, 15:50:13
Job time : 11.2651 secs

XX Claim 25; SEQ ID NO 6; 78pp; English.
XX The present invention relates to a hybrid protein consisting essentially
CC of the fusion of a membrane protein with an ion channel which is not
CC naturally coupled to the membrane protein. The hybrid protein is used as
CC an electrical sensor for screening of an agonist/antagonist of a membrane
CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel Kir6.2 mutant (KR370AA) protein.
XX
SQ Sequence 1927 AA;
Query Match 100.0%; Score 9901; DB 9; Length 1927;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1927; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSADGSDPLMDNMVNTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFPLYLSRH 60
DB 1 MALRGFCSADGSDPLMDNMVNTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFPLYLSRH 60
QY 61 DRGYIOMTPLNKKTALGFLIMVCMADLFYSFMRSGIFLAPVFLVSPILLGITTLA 120
DB 61 DRGYIOMTPLNKKTALGFLIMVCMADLFYSFMRSGIFLAPVFLVSPILLGITTLA 120
QY 121 TELLQLERRKGVSSGIMLTFMVLVLCALAIIRSKIMTALKEPAQVDLPFDIFPYVES 180
DB 121 TELLQLERRKGVSSGIMLTFMVLVLCALAIIRSKIMTALKEPAQVDLPFDIFPYVES 180
QY 181 LLLIQVLVSCFSDSPFLSETIHDNPPCESASFLSRTFMWITGLIVRGYROPLEBSD 240
DB 181 LLLIQVLVSCFSDSPFLSETIHDNPPCESASFLSRTFMWITGLIVRGYROPLEBSD 240
QY 241 LMSLNKEDTSQVVPVLYVNMKKCECAKTRKQPVKVYSSKDPAQKESKTDANEVVAL 300
DB 241 LMSLNKEDTSQVVPVLYVNMKKCECAKTRKQPVKVYSSKDPAQKESKTDANEVVAL 300
QY 301 IVKSPQKEMNPSLFLVLYKTFGRPYLMSFFPKAIHDLMMFSGPOLKLLIKPVNDTKAPD 360
DB 301 IVKSPQKEMNPSLFLVLYKTFGRPYLMSFFPKAIHDLMMFSGPOLKLLIKPVNDTKAPD 360
QY 361 WQGFYFVTLFVTAQLQTLVLYHOYFHIQFVSGMRIKTAVIGAVYKALVITNSAKSSTV 420
DB 361 WQGFYFVTLFVTAQLQTLVLYHOYFHIQFVSGMRIKTAVIGAVYKALVITNSAKSSTV 420
QY 421 GEIYNLMSVDAQRFMDLATYINMTWSAPLOVILALYLLMLNGPSVLAGVAVMLMEVFN 480
DB 421 GEIYNLMSVDAQRFMDLATYINMTWSAPLOVILALYLLMLNGPSVLAGVAVMLMEVFN 480
QY 481 AVMMKTKTYOAVAHKSKDNRIKLMNELINGIKVLYAMELAFKDYLAIRQEBLKYLK 540
DB 481 AVMMKTKTYOAVAHKSKDNRIKLMNELINGIKVLYAMELAFKDYLAIRQEBLKYLK 540
QY 541 KSAVLSAVGTFWVCTPFLVALCTPAVYVTTIDENNILDQAQFVSLAFNILRFLANTLP 600
DB 541 KSAVLSAVGTFWVCTPFLVALCTPAVYVTTIDENNILDQAQFVSLAFNILRFLANTLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEELPDSIERRPVKDGGGTNSITVRNATFTWASDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEELPDSIERRPVKDGGGTNSITVRNATFTWASDPT 660
QY 661 LNGITFSPREGALVAVGVQVGGKSSLSALLAEMDKVEGHAIRGSVAVYPQQAQWIOND 720
DB 661 LNGITFSPREGALVAVGVQVGGKSSLSALLAEMDKVEGHAIRGSVAVYPQQAQWIOND 720
QY 721 SLRENIILGCOLLEBYRSVIOACALPLDLTLPDGRTEIGEKVNSGGOKQVSLAR 780
DB 721 SLRENIILGCOLLEBYRSVIOACALPLDLTLPDGRTEIGEKVNSGGOKQVSLAR 780
QY 781 AVYSNADIYLPDDPLSAVDAHVGHKIPENVIGPKMKLNKTRILVTHSMSTYLPQVDVITV 840
DB 781 AVYSNADIYLPDDPLSAVDAHVGHKIPENVIGPKMKLNKTRILVTHSMSTYLPQVDVITV 840

DB 781 AVYSNADIYLPDDPLSAVDAHVGHKIPENVIGPKMKLNKTRILVTHSMSTYLPQVDVITV 840
QY 841 MSGGKISEMSYQELIARDGAFAEFLRTVASTEOQDAENGVTGVSQPGKAKOMENGM 900
DB 841 MSGGKISEMSYQELIARDGAFAEFLRTVASTEOQDAENGVTGVSQPGKAKOMENGM 900
QY 901 LVYDSACKQLOROLSSSSSYSGDISRRHNSTAEIOKAEMKEETWKIMEDAKOTGOVKL 960
DB 901 LVYDSACKQLOROLSSSSSYSGDISRRHNSTAEIOKAEMKEETWKIMEDAKOTGOVKL 960
QY 961 SVYWDYKKAIGLFISFLISFLFMCNHSALASNTWLSLMTDDPVTNQTQEHTRLSVYG 1020
DB 961 SVYWDYKKAIGLFISFLISFLFMCNHSALASNTWLSLMTDDPVTNQTQEHTRLSVYG 1020
QY 1021 ALGISQGIADVGYMAVSIIGIILASRCLHYDLIHSIRPMSFPERTPSGNLVNRFSEKL 1080
DB 1021 ALGISQGIADVGYMAVSIIGIILASRCLHYDLIHSIRPMSFPERTPSGNLVNRFSEKL 1080
QY 1081 DTVDSMIPEVIKMGSLFNVIGACIYILLATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
DB 1081 DTVDSMIPEVIKMGSLFNVIGACIYILLATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
QY 1141 KRLSVSRSPVYSHFNFTLGVSVYIRAEQERFIHOSDLKVDENQAKYPSIVANRWLA 1200
DB 1141 KRLSVSRSPVYSHFNFTLGVSVYIRAEQERFIHOSDLKVDENQAKYPSIVANRWLA 1200
QY 1201 VRLCEVGNCIYLFALPAVISRHSLSAGLVGLSVYSLOVTTYLNMVLRMSSEMETNIVA 1260
DB 1201 VRLCEVGNCIYLFALPAVISRHSLSAGLVGLSVYSLOVTTYLNMVLRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVRVEFRNYCLRYREDLPVLNHNVTINGG 1320
DB 1261 VERLKEYSETEKEAPWQIOETAPSSWPQVRVEFRNYCLRYREDLPVLNHNVTINGG 1320
QY 1321 EKVGIQRTGAGKSLTGLFRINESAGEIIDIINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIQRTGAGKSLTGLFRINESAGEIIDIINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGSLRMNLDPFSSQYSDSEVWTSLELAHLKDFVSAHPKLDHECAEGENLSVGOQOLVCL 1440
DB 1381 SGSLRMNLDPFSSQYSDSEVWTSLELAHLKDFVSAHPKLDHECAEGENLSVGOQOLVCL 1440
QY 1441 ARALLRKTILVIDEATPAVDLETTDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500
DB 1441 ARALLRKTILVIDEATPAVDLETTDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIL 1500
QY 1501 DKGEIOEYGAESDILQORGLFYSAKADAGLVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
DB 1501 DKGEIOEYGAESDILQORGLFYSAKADAGLVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
QY 1561 PRYRTERBARFVSKKGCNVAHKNIREQGRFLQDVFTTVLDLKMPTLLIFTMSFLCSW 1620
DB 1561 PRYRTERBARFVSKKGCNVAHKNIREQGRFLQDVFTTVLDLKMPTLLIFTMSFLCSW 1620
QY 1621 LLFAMVWMLIAFAHGDLAPGEGTVPCVTSIHSFSAFLFSIBVQVTIGFGGRWTEBCP 1680
DB 1621 LLFAMVWMLIAFAHGDLAPGEGTVPCVTSIHSFSAFLFSIBVQVTIGFGGRWTEBCP 1680
QY 1681 LAIILLYQNVGMINAMIGCIPMKTQAQHRRAETLIFSKHAVITLIRHRLCEPMARVG 1740
DB 1681 LAIILLYQNVGMINAMIGCIPMKTQAQHRRAETLIFSKHAVITLIRHRLCEPMARVG 1740
QY 1741 DLKRSMTISATIHQVARKTSSPEGEVYPLHQVDIPMENGVGNGSIFLVALIITHVDS 1800
DB 1741 DLKRSMTISATIHQVARKTSSPEGEVYPLHQVDIPMENGVGNGSIFLVALIITHVDS 1800
QY 1801 NSPLYDLAPSDLIHHQDLEIIVILEGVEITGITTOARTSYLADEIIMGQRFVPIVAED 1860
DB 1801 NSPLYDLAPSDLIHHQDLEIIVILEGVEITGITTOARTSYLADEIIMGQRFVPIVAED 1860
QY 1861 GRYSVDYSKFGNTIKVPTPLCTARQULDEDRSLDALTLASSRGLRAASVAVAKAKPFS 1920
DB 1861 GRYSVDYSKFGNTIKVPTPLCTARQULDEDRSLDALTLASSRGLRAASVAVAKAKPFS 1920

QY 1921 ISPDLS 1927
Db 1921 ISPDLS 1927

RESULT 2
ID ADY86938 standard; protein. 1927 AA.
AC ADY86938;
DT 02-JUN-2005 (first entry)
XX
DE Human MRP1-mouse Klr6.2 fusion protein, SEQ ID NO: 1.
XX
KM Ionophore; biosensor; drug screening; diagnostic;
KW microorganism detection; potassium channel; fusion protein;
XX multidrug resistance protein 1; MRP1, Klr6.2.
OS Homo sapiens.
OS Mus musculus.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Region 1..1531
FT /note= "Multidrug resistance protein 1 (MRP1)"
FT Region 1532..1537
FT /note= "Hexaglycine spacer"
FT Region 1538..1927
FT /note= "Klr6.2 protein"

XX US2005063989-A1.
XX
XX 24-MAR-2005.
XX
XX 22-SEP-2003; 2003US-00665283.
XX
XX 22-SEP-2003; 2003US-00665283.
XX
XX (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
XX
XX Derand R, Garcia E, Prost A, Revilloud J, Vivaudou M,
XX WPI; 2005-252611/26.
XX
XX
XX New ion channel hybrid protein, used as electrical sensor for screening
PT an agonist/antagonist of a membrane protein and for detecting a
PT contaminant/pollutant in a sample.
XX
XX
XX Claim 25; SEQ ID NO 1; 78bp; English.
XX
XX The present invention relates to a hybrid protein consisting essentially
CC of the fusion of a membrane protein with an ion channel which is not
CC naturally coupled to the membrane protein. The hybrid protein is used as
CC an electrical sensor for screening of an agonist/antagonist of a membrane
CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel Klr6.2 protein.
XX
XX
SQ Sequence 1927 AA;

Query Match 99.9%; Score 9891; DB 9; Length 1927;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1925; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MARGFSGADGSPILMNMNTWTNSNDPTKCFONTYLVWPPCFYLMACFPFYLYLSRH 60
Db 1 MARGFSGADGSPILMNMNTWTNSNDPTKCFONTYLVWPPCFYLMACFPFYLYLSRH 60

QY 61 DRGYIQWTPLNKTALGFLIMIVCMADLFYSFWRSGRGIPLAVFLVSPTLGITTLLA 120

Db 61 DRGYIQWTPLNKTALGFLIMIVCMADLFYSFWRSGRGIPLAVFLVSPTLGITTLLA 120
QY 121 TFLIQERRRGVSSGIMLTFWLVALVCAIAIIRSKITMTALKEDAQVDFRDIITYYFS 180
Db 121 TFLIQERRRGVSSGIMLTFWLVALVCAIAIIRSKITMTALKEDAQVDFRDIITYYFS 180
QY 181 LLLIQVLSCPSRSLPSETIHDNCPRESSASFLSRIFEMWTGILVYGQPLGSD 240
Db 181 LLLIQVLSCPSRSLPSETIHDNCPRESSASFLSRIFEMWTGILVYGQPLGSD 240
QY 241 LMSLNKEDTSEQVVPVVKMKKECAKTRKQPVVYSSKDPQPKSSKVDABEVEAL 300
Db 241 LMSLNKEDTSEQVVPVVKMKKECAKTRKQPVVYSSKDPQPKSSKVDABEVEAL 300
QY 301 IVSPQKEMNPSLFKVLKTFGGPYFLMSFPFKAIHIDLMFSGPQILKLIKPVNDYAPD 360
Db 301 IVSPQKEMNPSLFKVLKTFGGPYFLMSFPFKAIHIDLMFSGPQILKLIKPVNDYAPD 360
QY 361 WQGYFTVLLFVTAQCLQTLVHQYFHCFSVGRIRKTAIVGAYRKALVITNSARKSTV 420
Db 361 WQGYFTVLLFVTAQCLQTLVHQYFHCFSVGRIRKTAIVGAYRKALVITNSARKSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINMISAPLQVIALYLLMNLGSPVLAGVAVMVLAMYVNV 480
Db 421 GEIVNLSVDAQRFMDLATYINMISAPLQVIALYLLMNLGSPVLAGVAVMVLAMYVNV 480
QY 481 AVNARKTKTYQVAMHMSKDNRIKLMNEILNGIKVLKLYAMELAFKDKVLAIRQBELKVLK 540
Db 481 AVNARKTKTYQVAMHMSKDNRIKLMNEILNGIKVLKLYAMELAFKDKVLAIRQBELKVLK 540
QY 541 KSAVLSAVGFTFWCTPFLVALCTFAVYVITDENNILDQTAFLSALFNILRPDLNLP 600
Db 541 KSAVLSAVGFTFWCTPFLVALCTFAVYVITDENNILDQTAFLSALFNILRPDLNLP 600
QY 601 MWISIVQASVSLKRLIFLSHELEBPSIERRPVKDGGTNSITVNAATFTMARSPPT 660
Db 601 MWISIVQASVSLKRLIFLSHELEBPSIERRPVKDGGTNSITVNAATFTMARSPPT 660
QY 661 LMGITTSIPBGALVAVVGQVCGKSSLSALALBEMDKVEGHVAIKGSVAVVPQAWIOND 720
Db 661 LMGITTSIPBGALVAVVGQVCGKSSLSALALBEMDKVEGHVAIKGSVAVVPQAWIOND 720
QY 721 SLRENILFGQLEBPYRSYTOACALLPDLIELIPSGDRTIIGKGVNLSCGQKORVSLAR 780
Db 721 SLRENILFGQLEBPYRSYTOACALLPDLIELIPSGDRTIIGKGVNLSCGQKORVSLAR 780
QY 781 AVYSNADIVLPDPLASVDHVGKHFENVYIGPKMLKNTRIILVTHSMGYLPQVDYIIV 840
Db 781 AVYSNADIVLPDPLASVDHVGKHFENVYIGPKMLKNTRIILVTHSMGYLPQVDYIIV 840
QY 841 MSGGKISEMGSYOELLARDGAPAEFLRTYASTEOBDAENGTVGSGPKKAKOMENGM 900
Db 841 MSGGKISEMGSYOELLARDGAPAEFLRTYASTEOBDAENGTVGSGPKKAKOMENGM 900
QY 901 LVTDSAGKQLOROLSSSSSYSGDISRHNSSTAELQRAEAKKETWKLMEADKAQTQGYKL 960
Db 901 LVTDSAGKQLOROLSSSSSYSGDISRHNSSTAELQRAEAKKETWKLMEADKAQTQGYKL 960
QY 961 SVYWDYWKALGLISFLSTFLFMCNHYSAIASYVWISLWTDDEILVNGTOSHTRYLSVYG 1020
Db 961 SVYWDYWKALGLISFLSTFLFMCNHYSAIASYVWISLWTDDEILVNGTOSHTRYLSVYG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLHVDLHSLIRSPMSFERTPSGNLVNRFSEKL 1080
Db 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLHVDLHSLIRSPMSFERTPSGNLVNRFSEKL 1080
QY 1081 DTVDMSIPEYIKNPMGSLFNVIGACIVILLATPIAIIIPPLGLIYFVQRFVASSROL 1140
Db 1081 DTVDMSIPEYIKNPMGSLFNVIGACIVILLATPIAIIIPPLGLIYFVQRFVASSROL 1140
QY 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEEQEFHQSQDKVDENOKAYPSIVANRWLA 1200

```

Db      1141 KRLESVRSRSPVYSHFNETLLGVSVIRAEFEQERFIHQSDLKVDENQKAYPSIVANRMLA 1200
Qy      1201 VRLECVGNCLVLFALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETINVA 1260
Db      1201 VRLECVGNCLVLFALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETINVA 1260
Qy      1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVEFRNYCLARYREDDLVLRHINVTINGG 1320
Db      1261 VERLKEYSETEKEAPWQIQETAPSSWPQVGRVEFRNYCLARYREDDLVLRHINVTINGG 1320
Qy      1321 EKVGIVERTGAGKSLTGLFRINESAGEIITIGINIAKIGLHDLRKTITITIDPPLF 1380
Db      1321 EKVGIVERTGAGKSLTGLFRINESAGEIITIGINIAKIGLHDLRKTITITIDPPLF 1380
Qy      1381 SGSLRMLNDPFSQYSDSEEWMTSLBLAHLDKDVSAIPDLDBHECAGGENTLVSQROLYCL 1440
Db      1381 SGSLRMLNDPFSQYSDSEEWMTSLBLAHLDKDVSAIPDLDBHECAGGENTLVSQROLYCL 1440
Qy      1441 ARALLRKTIIIVDEBATAVDLETDLLIQSTIRFQEDCTVLTIAHRLNTIMDYTRIVL 1500
Db      1441 ARALLRKTIIIVDEBATAVDLETDLLIQSTIRFQEDCTVLTIAHRLNTIMDYTRIVL 1500
Qy      1501 DKGEIOEYAGSDLLQQRGLFYSAKXDGIVGGGGMLSRKGIIPREYVLTIRLAEDPAE 1560
Db      1501 DKGEIOEYAGSDLLQQRGLFYSAKXDGIVGGGGMLSRKGIIPREYVLTIRLAEDPAE 1560
Qy      1561 PRYTRERRARFVSKGKCNVAHKNIREQGRFLQDVFTTIVDLKWPHTLLIFTSFCLSM 1620
Db      1561 PRYTRERRARFVSKGKCNVAHKNIREQGRFLQDVFTTIVDLKWPHTLLIFTSFCLSM 1620
Qy      1621 LLFAMWMLIAFAHGDLAGEGTNVPCVTSIHSFSSAFLEFISIEVOVTIGFGGRWTEBCP 1680
Db      1621 LLFAMWMLIAFAHGDLAGEGTNVPCVTSIHSFSSAFLEFISIEVOVTIGFGGRWTEBCP 1680
Qy      1681 LATILIVQNVGLMINAIMGCTFMKTAQAHRAEETLIFSKHANTIRHRLCQMLRAG 1740
Db      1681 LATILIVQNVGLMINAIMGCTFMKTAQAHRAEETLIFSKHANTIRHRLCQMLRAG 1740
Qy      1741 DLKSMIISATIHQOVAKTSPSEGEVPLHCVDPFMENGVGANGIFLVAPLIYHVIDS 1800
Db      1741 DLKSMIISATIHQOVAKTSPSEGEVPLHCVDPFMENGVGANGIFLVAPLIYHVIDS 1800
Qy      1801 NSPIYDLAPSDLHHQDLEIIVLEGVETGITTQARTSYLADELWGQRFVPIVAED 1860
Db      1801 NSPIYDLAPSDLHHQDLEIIVLEGVETGITTQARTSYLADELWGQRFVPIVAED 1860
Qy      1861 GRYSVDYKFGNTIKVPTPLCTARQLDDBDRSLDLTLASSRGPILRAASVAVAKKPKFS 1920
Db      1861 GRYSVDYKFGNTIKVPTPLCTARQLDDBDRSLDLTLASSRGPILRAASVAVAKKPKFS 1920
Qy      1921 ISPDSL 1927
Db      1921 ISPDSL 1927

```

RESULT 3
ADY86945
ID ADY86945 standard; protein; 1947 AA.

XX AC ADY86945;
XX DT 02-JUN-2005 (first entry)
XX DE Human MRP1-mouse Kir6.2-HA fusion protein, SEQ ID NO: 8.
XX KW Ionophore; biosensor; drug screening; diagnostic;
XX KW microorganism detection; potassium channel; fusion protein;
XX KW multidrug resistance protein 1; MRP1; Kir6.2.
XX OS Homo sapiens.
XX OS Mus musculus.
XX OS Chimeric.
OS Unidentified.

```

XX Key Location/Qualifiers
FH Region 1..1531
FT /note= "Multidrug resistance protein 1 (MRP1)"
FT Region 1532..1537
FT /note= "Hexaglycine spacer"
FT Region 1538..1947
FT /note= "Kir6.2 protein containing HA epitope peptide"
FT Region 1651..1659
FT /note= "HA epitope peptide"
XX US2005063989-A1.
XX 24-MAR-2005.
XX 22-SEP-2003; 2003US-00665283.
XX 22-SEP-2003; 2003US-00665283.
XX (COMS ) COMMISSARIAT ENERGIE ATOMIQUE.
PI Derand R, Garcia E, Prost A, Revilloud J, Vivaudou M,
XX WPI; 2005-252611/26.
XX
XX New ion channel hybrid protein, used as electrical sensor for screening
PT an agonist/antagonist of a membrane protein and for detecting a
XX contaminant/pollutant in a sample.
PS Claim 25; SEQ ID NO 8; 78pp; English.
XX
XX The present invention relates to a hybrid protein consisting essentially
CC of the fusion of a membrane protein with an ion channel which is not
CC naturally coupled to the membrane protein. The hybrid protein is used as
CC an electrical sensor for screening of an agonist/antagonist of a membrane
CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel Kir6.2 protein containing HA epitope peptide.
XX
XX Sequence 1947 AA:
SQ
Query Match 99.6%; Score 9861; DB 9; Length 1947;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1925; Conservative 0; Mismatches 2; Indels 20; Gaps 2;
Qy 1 MALRGFSGADGSDPLMDMNTWTWNTSNPDFTKCFQNTVLLVWPCTYLMACFPFFLYLSRH 60
Db 1 MALRGFSGADGSDPLMDMNTWTWNTSNPDFTKCFQNTVLLVWPCTYLMACFPFFLYLSRH 60
Qy 61 DRGIQMTPLNKTATGFLIMIVCWADLFYSPWERSRGIFLAVPLVSPTLGTTTLA 120
Db 61 DRGIQMTPLNKTATGFLIMIVCWADLFYSPWERSRGIFLAVPLVSPTLGTTTLA 120
Qy 121 TFLIQLEBRKGVQSSGIMLTFMVALVCALALRSKIMTALKEDAOVDLFRDITFYVYS 180
Db 121 TFLIQLEBRKGVQSSGIMLTFMVALVCALALRSKIMTALKEDAOVDLFRDITFYVYS 180
Qy 181 LLLIQVLVISCPSDRSPLESETIHDNPNCPBSSASFLSRTITWMTTGLIVRGROPLBESD 240
Db 181 LLLIQVLVISCPSDRSPLESETIHDNPNCPBSSASFLSRTITWMTTGLIVRGROPLBESD 240
Qy 241 LMSLNKEDTSROVVPVLYKMKKCKAKRKQPVVVVSSKQPAQKSSKVDANEVEVAL 300
Db 241 LMSLNKEDTSROVVPVLYKMKKCKAKRKQPVVVVSSKQPAQKSSKVDANEVEVAL 300
Qy 301 IVKSPQKEMNSLPKVLVYKTFGPFYLMFFPKAHLDMFMFGPOLKLLIFVNDTKAPD 360
Db 301 IVKSPQKEMNSLPKVLVYKTFGPFYLMFFPKAHLDMFMFGPOLKLLIFVNDTKAPD 360
Qy 361 WQGYFTVLLFVTACLOTVLHQYFHI CFVSGMRITKAVIGAVYRKALVITNSARKSSTV 420

```

Db 361 WQGYFTVLLFVTAQCTLVHGHQFHLICFVSGMRITKFAVIGVYRKALVITNSARKSTV 420
 Qy 421 GEIYNLMSVDAQRPMDLATYTNMTMSAPLOYITLALYLMNLGSSVLAGVAVMTLMPVN 480
 Db 421 GEIYNLMSVDAQRPMDLATYTNMTMSAPLOYITLALYLMNLGSSVLAGVAVMTLMPVN 480
 Qy 481 AVMAKTKTYOVAMHMSKDNRIKLMNEILNGIKYVKLYAMELAFDKYLAIOBELKYLK 540
 Db 481 AVMAKTKTYOVAMHMSKDNRIKLMNEILNGIKYVKLYAMELAFDKYLAIOBELKYLK 540
 Qy 541 KSAVLSAVGTFTWCTPPLVALCTPAVYVTIDENNIIIDAQTAFLSLFNILREPLNLP 600
 Db 541 KSAVLSAVGTFTWCTPPLVALCTPAVYVTIDENNIIIDAQTAFLSLFNILREPLNLP 600
 Qy 601 MVISSIVQASVSLRLRIFLSHELEPDSIRRPVYKGGGNTSTTNNATTTMARSDPT 660
 Db 601 MVISSIVQASVSLRLRIFLSHELEPDSIRRPVYKGGGNTSTTNNATTTMARSDPT 660
 Qy 661 LMGITFSPREGALVAVVGVCGCKSSLLSALLAEMDKVGHVATKGSVAVVPOQAMIOND 720
 Db 661 LMGITFSPREGALVAVVGVCGCKSSLLSALLAEMDKVGHVATKGSVAVVPOQAMIOND 720
 Qy 721 SLRENILFGCQLEBPYRSVIOACALLPDLIELPSGDRTEIGEGVNLSSGQKQVSLAR 780
 Db 721 SLRENILFGCQLEBPYRSVIOACALLPDLIELPSGDRTEIGEGVNLSSGQKQVSLAR 780
 Qy 781 AVYSNADITYLFDDELAVDAHVGHIFBNVIGPYGMKNKTRILVTHSMSTLPQVDYIIV 840
 Db 781 AVYSNADITYLFDDELAVDAHVGHIFBNVIGPYGMKNKTRILVTHSMSTLPQVDYIIV 840
 Qy 841 MSGGKISEMGSYOELRLARDGFAFEFLRTYASTEOEADAENGVTGVSQPGKEAOMENGM 900
 Db 841 MSGGKISEMGSYOELRLARDGFAFEFLRTYASTEOEADAENGVTGVSQPGKEAOMENGM 900
 Qy 901 LVTDASGKOLQROUSSSSYSGDISRRHNSTAEIOKAKEKEFTMKLEADKAOTGOVKL 960
 Db 901 LVTDASGKOLQROUSSSSYSGDISRRHNSTAEIOKAKEKEFTMKLEADKAOTGOVKL 960
 Qy 961 SVYWDYMKAIGLFTSFLIFMGNHVSALASNTWLSLMTDDPIVNGTOEHTKYRLSVYG 1020
 Db 961 SVYWDYMKAIGLFTSFLIFMGNHVSALASNTWLSLMTDDPIVNGTOEHTKYRLSVYG 1020
 Qy 1021 ALGISOGIAVFGVMAVSIIGIILASRCLHVDLHSHIIRSPSPFEERPSGULVRFSEKL 1080
 Db 1021 ALGISOGIAVFGVMAVSIIGIILASRCLHVDLHSHIIRSPSPFEERPSGULVRFSEKL 1080
 Qy 1081 DTVDSMTPEVYIKMFMGSLFNVIAGACIVILATPIAIIIPPLGLIFPFQRFYVASSROL 1140
 Db 1081 DTVDSMTPEVYIKMFMGSLFNVIAGACIVILATPIAIIIPPLGLIFPFQRFYVASSROL 1140
 Qy 1141 KRLSVSRSPYSHFNETLLGVSVIRAFEOERFIHOSDLKVDENQKAYPSIVANRWLA 1200
 Db 1141 KRLSVSRSPYSHFNETLLGVSVIRAFEOERFIHOSDLKVDENQKAYPSIVANRWLA 1200
 Qy 1201 VRLCEVNCNCLVFPALFAVISRHSLSAGVGLSVSYLQVTTYINMLVRMSSEMETIVA 1260
 Db 1201 VRLCEVNCNCLVFPALFAVISRHSLSAGVGLSVSYLQVTTYINMLVRMSSEMETIVA 1260
 Qy 1261 VERLKEVSETEKAPMOIOETAPSSWPQVGRVFRNYCJRYREDDLPVLRIHINTINGG 1320
 Db 1261 VERLKEVSETEKAPMOIOETAPSSWPQVGRVFRNYCJRYREDDLPVLRIHINTINGG 1320
 Qy 1321 EKVGIVRTGAGKSSLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
 Db 1321 EKVGIVRTGAGKSSLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
 Qy 1381 SSGLRNMLDPPSOYSDDEWVTSLELAHKDFVSALPKLHECABEGENISVGOROIYCL 1440
 Db 1381 SSGLRNMLDPPSOYSDDEWVTSLELAHKDFVSALPKLHECABEGENISVGOROIYCL 1440
 Qy 1441 ARALLRKTJLVLDEATAVADLETDDLIOSTIRTOFEDCVTLTAHRLANTIMDTRVIVL 1500
 Db 1441 ARALLRKTJLVLDEATAVADLETDDLIOSTIRTOFEDCVTLTAHRLANTIMDTRVIVL 1500

Qy 1501 DKGEIOEYGA PSDLLQORGLFYSMADAGLVGGGGGMLSRKGIIPREYVLTURLAEDPAE 1560
 Db 1501 DKGEIOEYGA PSDLLQORGLFYSMADAGLVGGGGGMLSRKGIIPREYVLTURLAEDPAE 1560
 Qy 1561 PRYTRERRRARFVSKKNCNVAAHKNIREOGRFLQDVFETTLVDLKMPTLLIFTMSFLCSW 1620
 Db 1561 PRYTRERRRARFVSKKNCNVAAHKNIREOGRFLQDVFETTLVDLKMPTLLIFTMSFLCSW 1620
 Qy 1621 LLEFAMVWMLIAFAHG-----DLAP-----GEGTNVPCVTSIHSFSSAFLF 1680
 Db 1621 LLEFAMVWMLIAFAHG-----DLAP-----GEGTNVPCVTSIHSFSSAFLF 1680
 Qy 1681 SIEVQVITIGFGRVNTSECPALILIVONIVGIMINAMLCIFMKTAAHRAEFLIF 1740
 Db 1681 SIEVQVITIGFGRVNTSECPALILIVONIVGIMINAMLCIFMKTAAHRAEFLIF 1740
 Qy 1721 SKHAVITLRRGRCLCFMLRVGDLRKSMIISATIHQVVRKTTSPGEVVPPLHVDIIPMBNG 1780
 Db 1721 SKHAVITLRRGRCLCFMLRVGDLRKSMIISATIHQVVRKTTSPGEVVPPLHVDIIPMBNG 1780
 Qy 1781 VGGNGIFLVAPLIIVHVIDNSPLYDLAPSDLHHODLEIIVILEGVETTGITTOARTS 1840
 Db 1801 VGGNGIFLVAPLIIVHVIDNSPLYDLAPSDLHHODLEIIVILEGVETTGITTOARTS 1860
 Qy 1841 YLADEILMGORFPPIVAEBDGRYSVDYSKRGNTIKVPTPLCTARQLDEDSLLDALTLAS 1900
 Db 1861 YLADEILMGORFPPIVAEBDGRYSVDYSKRGNTIKVPTPLCTARQLDEDSLLDALTLAS 1920
 Qy 1901 SRGPLRAASVAVAKAPKPSISPDLS 1927
 Db 1921 SRGPLRAASVAVAKAPKPSISPDLS 1947
 RESULT 4
 ID ADY86941 standard; protein; 1891 AA.
 XX ADY86941;
 AC XX
 DT 02-JUN-2005 (first entry)
 XX
 DE Human MRP1-Kir6.2 mutant delac36 fusion protein, SEQ ID NO: 4.
 XX
 KW Ionophore; biosensor; drug screening; diagnostic;
 KM microorganism detection; potassium channel; fusion protein;
 XX multidrug resistance protein 1; MRP1; Kir6.2; mutein.
 OS Homo sapiens.
 OS Mus musculus.
 OS Chimeric.
 XX
 FT Region location/Qualifiers
 FT 1..1531
 FT /note="Multidrug resistance protein 1 (MRP1)"
 FT Region 1532..1537
 FT /note="Hexaglycine spacer"
 FT 1538..1891
 FT /note="Kir6.2 mutant delac36 protein"
 XX
 PN US2005063989-A1.
 XX
 PD 24-MAR-2005.
 XX
 PF 22-SEP-2003; 2003US-00665283.
 XX
 PR 22-SEP-2003; 2003US-00665283.
 XX
 PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
 XX
 PI Derand R, Garcia E, Prost A, Revilloud J, Vivaudou M;
 XX WPI; 2005-252611/26.

XX New ion channel hybrid protein, used as electrical sensor for screening
PT an agonist/antagonist of a membrane protein and for detecting a
PT contaminant/pollutant in a sample.
XX
PS Claim 25; SEQ ID NO 4; 78bp; English.
XX
CC The present invention relates to a hybrid protein consisting essentially
CC of the fusion of a membrane protein with an ion channel which is not
CC naturally coupled to the membrane protein. The hybrid protein is used as
CC an electrical sensor for screening of an agonist/antagonist of a membrane
CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel Kir6.2 mutant detac36 protein.
XX
SQ Sequence 1891 AA.
Query Match 98.3%; Score 9734; DB 9; Length 1891;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALPGCSADGSDPLMDMNTWNTSNPDFTKCFQNTVLWVPCFLYLMACFPFFLYLSRH 60
DB 1 MALPGCSADGSDPLMDMNTWNTSNPDFTKCFQNTVLWVPCFLYLMACFPFFLYLSRH 60
QY 61 DRGYIOMTPLNKTALGFLIMIVCMADLFYSFMRSGIFLAPFLVSPILGITTLA 120
DB 61 DRGYIOMTPLNKTALGFLIMIVCMADLFYSFMRSGIFLAPFLVSPILGITTLA 120
QY 121 TFLIQLEBRKGVSGGIMLTFWLVALYCALAIRSKIMTALKEDAQVDLPFDITFYVYS 180
DB 121 TFLIQLEBRKGVSGGIMLTFWLVALYCALAIRSKIMTALKEDAQVDLPFDITFYVYS 180
QY 181 LLLIQVLVSCSDSPLEFSETHDNPCESSASLSTITFWMTGLIVRGYRQPLESD 240
DB 181 LLLIQVLVSCSDSPLEFSETHDNPCESSASLSTITFWMTGLIVRGYRQPLESD 240
QY 181 LLLIQVLVSCSDSPLEFSETHDNPCESSASLSTITFWMTGLIVRGYRQPLESD 240
DB 181 LLLIQVLVSCSDSPLEFSETHDNPCESSASLSTITFWMTGLIVRGYRQPLESD 240
QY 241 LMSLINKEDTSBOQVPLVYKMKKECAKTRKOPVKVYVSSKDPADPKESSKVDANEVVAL 300
DB 241 LMSLINKEDTSBOQVPLVYKMKKECAKTRKOPVKVYVSSKDPADPKESSKVDANEVVAL 300
QY 301 IVKSPQKEMNPSLFKVLVYKTFGPYFLMSFPFKAIHDLMMFSGPOLIKLIRFVDTKAPD 360
DB 301 IVKSPQKEMNPSLFKVLVYKTFGPYFLMSFPFKAIHDLMMFSGPOLIKLIRFVDTKAPD 360
QY 361 WQGFYTYTLFVYTAQLQTLVHQYFHICFVSGMRKTAIVAGVYRKALVITNSARKSSTV 420
DB 361 WQGFYTYTLFVYTAQLQTLVHQYFHICFVSGMRKTAIVAGVYRKALVITNSARKSSTV 420
QY 421 GEIVNLMSVDAORFMDLATYINMISAPLOYILALYLLMLNGPSVLGAVMVLMPVYN 480
DB 421 GEIVNLMSVDAORFMDLATYINMISAPLOYILALYLLMLNGPSVLGAVMVLMPVYN 480
QY 481 AYMAKTKTYOVANHKSNDRIKLNNEILNGIKVLKLYAMELAFKDYALIROBELKYLK 540
DB 481 AYMAKTKTYOVANHKSNDRIKLNNEILNGIKVLKLYAMELAFKDYALIROBELKYLK 540
QY 541 KSAIYLSAVGTITWCTPTVALCTFAVYTTIDENNIIDAQTAFFUSLALFNILRPLNTLP 600
DB 541 KSAIYLSAVGTITWCTPTVALCTFAVYTTIDENNIIDAQTAFFUSLALFNILRPLNTLP 600
QY 541 KSAIYLSAVGTITWCTPTVALCTFAVYTTIDENNIIDAQTAFFUSLALFNILRPLNTLP 600
DB 541 KSAIYLSAVGTITWCTPTVALCTFAVYTTIDENNIIDAQTAFFUSLALFNILRPLNTLP 600
QY 601 MVISSIVQASVSLKRLIFLSHEBELPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
DB 601 MVISSIVQASVSLKRLIFLSHEBELPDSIERRPVKDGGTNSITVRNATFTWASDPT 660
QY 661 LAGTITFSIPGALVAVVGQVCGKSSLSALLAEMDKYEGHAIKGSVAYVPOQAMIOND 720
DB 661 LAGTITFSIPGALVAVVGQVCGKSSLSALLAEMDKYEGHAIKGSVAYVPOQAMIOND 720
QY 721 SLRENIIIFGCOLLEBPYRSVIOACALLPDLBITLPSGRTIEGKGVNLSSGQKORVSLAR 780
DB 721 SLRENIIIFGCOLLEBPYRSVIOACALLPDLBITLPSGRTIEGKGVNLSSGQKORVSLAR 780

DB 721 SLRENIIIFGCOLLEBPYRSVIOACALLPDLBITLPSGRTIEGKGVNLSSGQKORVSLAR 780
QY 781 AVYSNADIIYFDDPLSAVDAAVGHKIPENYIGPKMLKNKTRILVTHSMSTLPQVDVIY 840
DB 781 AVYSNADIIYFDDPLSAVDAAVGHKIPENYIGPKMLKNKTRILVTHSMSTLPQVDVIY 840
QY 841 MSGKISEMGSYQELIARDGAFAEFLTYASTEOBDAENGVTGVSQPKAKOMENG 900
DB 841 MSGKISEMGSYQELIARDGAFAEFLTYASTEOBDAENGVTGVSQPKAKOMENG 900
QY 901 LYVDSAKQIQRQULSSSSSYSGDLSRHNSIAELQKAEKKEETWKMEADKAGTGVKL 960
DB 901 LYVDSAKQIQRQULSSSSSYSGDLSRHNSIAELQKAEKKEETWKMEADKAGTGVKL 960
QY 961 SVYDWYKAKIGLFISFISIFLFMCNHSALASNYMLSLMTDDPIVNGTOEHTKRLSYG 1020
DB 961 SVYDWYKAKIGLFISFISIFLFMCNHSALASNYMLSLMTDDPIVNGTOEHTKRLSYG 1020
QY 1021 ALGISGIIAVFGYSMAVSIIGIILASRCLAYDLHSIIRSPMSFPERTPSGNLVNRFSEKL 1080
DB 1021 ALGISGIIAVFGYSMAVSIIGIILASRCLAYDLHSIIRSPMSFPERTPSGNLVNRFSEKL 1080
QY 1081 DTVDSMTPEVIXKPMGSLFNVIGACIVILATPIAIIIPPLGLIYFFVQRFYASSRQL 1140
DB 1081 DTVDSMTPEVIXKPMGSLFNVIGACIVILATPIAIIIPPLGLIYFFVQRFYASSRQL 1140
QY 1141 KRLSEVSRSPIYSHFNFTLGVSVIRAFEBQEREIHOSDKVDENOKAYPSIVANFWLA 1200
DB 1141 KRLSEVSRSPIYSHFNFTLGVSVIRAFEBQEREIHOSDKVDENOKAYPSIVANFWLA 1200
QY 1201 VRLCEVGNCTYLPAALPAVYSRHSLSAGLVGLSVYSLOVTTYVNLVRSSEMETNIVA 1260
DB 1201 VRLCEVGNCTYLPAALPAVYSRHSLSAGLVGLSVYSLOVTTYVNLVRSSEMETNIVA 1260
QY 1261 VERLKEYSERKEKPMQIOETAPSSSWPOVGRVFRNYCIRYREDLPVLRHINVTNGG 1320
DB 1261 VERLKEYSERKEKPMQIOETAPSSSWPOVGRVFRNYCIRYREDLPVLRHINVTNGG 1320
QY 1321 EKVGIVRTGAKSGLTLGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIVRTGAKSGLTLGLFRINESAGEIILIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGSLRMLNDPFSQYSDSEWVTSLELAHKDFVSAIPKLDHECAEGENLSVGORQVCL 1440
DB 1381 SGSLRMLNDPFSQYSDSEWVTSLELAHKDFVSAIPKLDHECAEGENLSVGORQVCL 1440
QY 1441 ARALLRKTILVDEATAVADLETDDLIOSTIRPOFEDCTVLTAAHRLNTIMDTRYVL 1500
DB 1441 ARALLRKTILVDEATAVADLETDDLIOSTIRPOFEDCTVLTAAHRLNTIMDTRYVL 1500
QY 1501 DKGEIOEYGAPSDILOQRGLFYSMAKDAGLVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
DB 1501 DKGEIOEYGAPSDILOQRGLFYSMAKDAGLVGGGGGMLSRKGIIPREYVLTRLAEDPAE 1560
QY 1561 PRYTRERRARFVSKKCNVAHKNIREQGFLODVFTTIVDLKWPHTLLIFTNSFLCSW 1620
DB 1561 PRYTRERRARFVSKKCNVAHKNIREQGFLODVFTTIVDLKWPHTLLIFTNSFLCSW 1620
QY 1621 LLFAMVWMLTAFAGDLPBGSTVPCVTSIHSSSAFLPSIEVOYTIIGREGRWTEBCEP 1680
DB 1621 LLFAMVWMLTAFAGDLPBGSTVPCVTSIHSSSAFLPSIEVOYTIIGREGRWTEBCEP 1680
QY 1681 LAIILIVONVGMINAMIGCFMKTQAQHRAEFLIRSKHAVITLRHRQLCFMLRVG 1740
DB 1681 LAIILIVONVGMINAMIGCFMKTQAQHRAEFLIRSKHAVITLRHRQLCFMLRVG 1740
QY 1741 DLKRSMTIASITIHQVVRKTSPEGEVVPPLHQVDIPNENGVGNGIFLVADPLLIYHVIDS 1800
DB 1741 DLKRSMTIASITIHQVVRKTSPEGEVVPPLHQVDIPNENGVGNGIFLVADPLLIYHVIDS 1800
QY 1801 NSPLYDLAPSDLHHHODELIIIVILEGVEETGITTOARTSYLADEIIMGQRFVPIVAED 1860
DB 1801 NSPLYDLAPSDLHHHODELIIIVILEGVEETGITTOARTSYLADEIIMGQRFVPIVAED 1860

OY	1861	GRISVDYSKFGNTIKVPTPLCTARQDDEDS	1891
Dd	1861	GRISVDYSKFGNTIKVPTPLCTARQDDEDS	1891
 RESULT 5			
ID	AAW57486	standard; protein; 1531 AA.	
AC	AAW57486;		
XX			
DT	14-AUG-1998	(first entry)	
DE			
XX			
DE	Human MRP variant ltpgpa (lei/pgpa).		
KM	Multidrug resistance-associated protein; MRP; tumour; human; variant;		
RW	multidrug resistance; MDR; leishmania P-glycoprotein; ltpgpa; lei/pgpa.		
XX			
OS	Homo sapiens.		
XX			
FT			
FT	Key	Location/Qualifiers	
FH	Misc-difference	685	/label= L685S
FT		/note= "wild-type Leu is replaced by Ser"	
FT	Misc-difference	1282	/label= R1282A
FT		/note= "wild-type Arg is replaced by Ala"	
XX			
PN	US5766880-A.		
PD	16-JUN-1998.		
PF			
PP	05-JUN-1995;	95US-00463092.	
XX			
PR	27-OCT-1992;	92US-00966923.	
PR	08-MAR-1993;	93US-00029340.	
PR	26-OCT-1993;	93US-00141893.	
PR	20-MAR-1995;	95US-00407207.	
PA	(TOOH) UNIV QUEEN'S KINGSTON.		
PI	Cole SP, Deeley RG;		
XX			
DR	WPJ: 1998-361687/31.		
XX	N-PDB; AAW31498.		
PT	DNA encoding protein associated with multi-drug resistance - useful for		
PT	as probe for identifying multi-drug resistant tumour cells.		
PS	Claim 1; Col 67-78; 82pp; English.		
XX			
CC	This represents a variant of the human multidrug resistance-associated		
CC	protein (MRP). This natural variant is a leishmania P-glycoprotein related		
CC	molecule ltpgpa (lei/pgpa). The human and murine MRP nucleic acid		
CC	molecules can be used as probes for identifying multidrug resistant		
CC	tumour cells by hybridisation to mRNA from tumour cells. The antisense		
CC	nucleic acid can be used to reverse multidrug resistance (MDR). A		
CC	recombinant expression vector containing the MRP nucleic acid molecules		
CC	operatively linked to at least one regulatory sequence can be used to		
CC	transform a host cell to produce a recombinant MDR-associated protein		
SO	Sequence 1531 AA:		
 Query Match 79.4%; Score 7860; DB 2; Length 1531; Best Local Similarity 100.0%; Pred. No. 0; Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
OY	1	MALRGFCAGDGSDPLMDNNVTWNSTSNPDTKCFONTVLVWVCFLIMACFPFYLSRH	60
Dd	1	MALRGFCAGDGSDPLMDNNVTWNSTSNPDTKCFONTVLVWVCFLIMACFPFYLYSH	60
OY	61	DGRIQTMTPLNKTATLGLFLLIWCADLFYFWERSRCIFLAPELVSPILLGITTTLA	120

D	b		61	D	R	G	V	I	Q	M	T	P	L	K	T	A	L	G	F	L	M	I	Y	C	M	A	D	L	F	S	F	M	E	R	S	G	I	F	L	A	P	V	L	S	P	T	L	L	G	T	T	L	A		120					
Q	y		121	T	F	L	I	Q	E	R	R	K	V	O	S	S	G	I	M	L	T	F	M	L	V	A	L	C	A	L	I	R	S	K	I	M	T	A	L	K	E	D	A	Q	V	D	L	F	R	D	I	T	E	Y	Y	F	S		180	
D	b		121	T	F	L	I	Q	E	R	R	K	V	O	S	S	G	I	M	L	T	F	M	L	V	A	L	C	A	L	I	R	S	K	I	M	T	A	L	K	E	D	A	Q	V	D	L	F	R	D	I	T	E	Y	Y	F	S		180	
Q	y		181	L	L	I	I	Q	L	V	L	S	C	S	D	S	P	L	F	S	E	T	I	H	D	N	P	C	P	E	S	S	A	F	L	S	I	T	T	M	T	I	G	L	I	V	R	G	R	O	P	L	E	G	S	D		240		
D	b		181	L	L	I	I	Q	L	V	L	S	C	S	D	S	P	L	F	S	E	T	I	H	D	N	P	C	P	E	S	S	A	F	L	S	I	T	T	M	T	I	G	L	I	V	R	G	R	O	P	L	E	G	S	D		240		
Q	y		241	L	M	S	I	N	K	E	D	S	O	V	P	V	L	V	K	M	K	K	C	A	K	R	K	O	P	V	K	V	Y	S	S	K	O	P	A	O	K	E	S	S	K	D	A	N	E	V	A	L		300						
D	b		241	L	M	S	I	N	K	E	D	S	O	V	P	V	L	V	K	M	K	K	C	A	K	R	K	O	P	V	K	V	Y	S	S	K	O	P	A	O	K	E	S	S	K	D	A	N	E	V	A	L		300						
Q	y		301	I	V	K	S	P	O	K	E	M	N	S	L	F	E	V	L	K	T	G	P	F	I	M	S	F	F	E	K	A	I	H	D	L	M	F	S	G	P	O	L	I	K	L	I	F	E	V	N	D	T	K	A	D		360		
D	b		301	I	V	K	S	P	O	K	E	M	N	S	L	F	E	V	L	K	T	G	P	F	I	M	S	F	F	E	K	A	I	H	D	L	M	F	S	G	P	O	L	I	K	L	I	F	E	V	N	D	T	K	A	D		360		
Q	y		361	W	O	G	F	E	Y	T	V	L	F	V	T	A	C	L	O	T	L	V	H	O	F	P	H	I	C	F	S	G	M	R	I	K	T	A	V	I	G	A	V	R	K	A	L	V	I	T	N	S	A	R	K	S	T	V		420
D	b		361	W	O	G	F	E	Y	T	V	L	F	V	T	A	C	L	O	T	L	V	H	O	F	P	H	I	C	F	S	G	M	R	I	K	T	A	V	I	G	A	V	R	K	A	L	V	I	T	N	S	A	R	K	S	T	V		420
Q	y		421	G	E	I	N	L	M	S	V	D	R	F	M	D	L	T	T	N	M	I	W	S	A	P	L	O	V	I	T	L	A	L	I	L	M	L	G	S	V	L	A	G	V	A	V	M	L	M	E	V	N		480					
D	b		421	G	E	I	N	L	M	S	V	D	R	F	M	D	L	T	T	N	M	I	W	S	A	P	L	O	V	I	T	L	A	L	I	L	M	L	G	S	V	L	A	G	V	A	V	M	L	M	E	V	N		480					
Q	y		481	A	V	M	A	K	T	T	Y	O	V	A	H	M	K	S	K	O	N	R	I	K	L	M	N	E	L	I	N	G	I	K	V																									

Db 1141 KRLSVSRSPVYSHFNETLGVSVIRAFEEQERFIHOSDLKVDENQAYPSIVANRMLA 1200
Qy 1201 VRLSECVNGCIYLFALFAVISRSHLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETNIVA 1260
Db 1201 VRLSECVNGCIYLFALFAVISRSHLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETNIVA 1260
Qy 1261 VERLKEYSETEKEAPWQIQETAPPSWPQVRVEFRNYCLARYRBDLDFLRIHINVTINGG 1320
Db 1261 VERLKEYSETEKEAPWQIQETAPPSWPQVRVEFRNYCLARYRBDLDFLRIHINVTINGG 1320
Qy 1321 EKVGIVERTGAGKSLTGLFRINESARGEIITIDGINIAKIGLHDLRKITITIDQDPLTF 1380
Db 1321 EKVGIVERTGAGKSLTGLFRINESARGEIITIDGINIAKIGLHDLRKITITIDQDPLTF 1380
Qy 1381 SGLSRMNLDPFSQYSDSEEWMTSLBLAHKDFVSALPDKLDHECAEGENLSVGQROLVCL 1440
Db 1381 SGLSRMNLDPFSQYSDSEEWMTSLBLAHKDFVSALPDKLDHECAEGENLSVGQROLVCL 1440
Qy 1441 ARALLRKTIIVLDBATPAVDLETDDLIQSTIRTOFEDCTVLTAAHRLNTMDYTRIVL 1500
Db 1441 ARALLRKTIIVLDBATPAVDLETDDLIQSTIRTOFEDCTVLTAAHRLNTMDYTRIVL 1500
Qy 1501 DKGEIOEYGAAPSDLLQQRGLFYMAKDGGLV 1531
Db 1501 DKGEIOEYGAAPSDLLQQRGLFYMAKDGGLV 1531

RESULT 6
AAW74471
ID AAW74471 standard; protein; 1531 AA.
AC AAW74471;
XX
XX 18-MAY-1999 (first entry)
DT
XX
XX Human multidrug resistance-associated protein variant.
DE
XX Multidrug resistance-associated protein; MDR; human; diagnosis;
KW MDR tumour cell identification; cancer therapy.
RV
XX
OS Homo sapiens.
XX
PN US5882875-A.
PD 16-MAR-1999.
XX
XX 05-JUN-1995; 95US-00462109.
PF
XX 27-OCT-1992; 92US-00966923.
PR 08-MAR-1993; 93US-00029340.
PR 26-OCT-1993; 93US-00141893.
PR 20-MAR-1995; 95US-00407207.
XX
PA (TOOH) UNIV QUEBENS KINGSTON.
XX
PI Cole SPC, Deeley RG;
XX
DR WPI: 1999-214061/18.
DR N-PSDB; AAX21977.
XX
XX Identifying a multidrug resistant tumor cell by contacting the cell with
PT an antibody/antigen-binding fragment - which binds to an expressed
PR protein encoded by multidrug resistance-associated protein (MRP) nucleic
PT acid.
XX
PS Claim 3; Col 69-80; 80pp; English.
XX
XX This sequence is the human multidrug resistance-associated (MDR) protein.
CC The invention relates to a method for identifying a multidrug resistant
CC (MDR) tumor cell. Compositions and methods utilizing the MDR proteins
CC can be used to treat patients with tumors displaying multidrug
CC resistance, particularly those displaying resistance to anthracyclines,
CC epipodophyllotoxins, vinca alkaloids, and hydrophobic drugs. The methods

CC for inhibiting/killing a MDR tumour cell can be useful for treating
CC breast cancer, leukaemias, fibrosarcomas, cervical cancer, gliomas,
CC thymomas, neuroblastomas and lung cancer. The MDR DNA sequences when
CC labeled are useful as molecular probes for diagnosing multidrug
CC resistance of a tumour (using cells from a tumour biopsy) and for
CC designing ribozymes which are capable of cleaving a single-stranded
CC nucleic acid encoding a protein having MRP activity. Recombinant
CC expression vectors containing human MDR coding sequences can be
CC transfected into a drug sensitive cell line to produce a protein in the
CC cell which confers MDR, protecting non-resistant non-tumour cells from
CC the effects of chemotherapeutics has major clinical importance. Cells
CC transformed with the MDR coding sequences are useful for testing
CC potential therapeutic agents for their effectiveness against MDR cells
CC and for identifying chemosensitizers of a therapeutic agent
CC
XX
SQ Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 2; Length 1531;
Best local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRGFCSADGSDPLMDMNVNTMTNSNDFTKCFQNTVLVWPVPCFYLWACPFYFLYLSRH 60
Db 1 MALRGFCSADGSDPLMDMNVNTMTNSNDFTKCFQNTVLVWPVPCFYLWACPFYFLYLSRH 60
Qy 61 DRGYIQTPLNKTKTALGFLIMVCMADLFYSFWERSRGIFLADVFLVSPTLGITTLLA 120
Db 61 DRGYIQTPLNKTKTALGFLIMVCMADLFYSFWERSRGIFLADVFLVSPTLGITTLLA 120
Qy 121 TFLIQLEBRKGVSSGIMLTFMLVALYCALAILASKIMTAKENAVDLPFDITFYVYFS 180
Db 121 TFLIQLEBRKGVSSGIMLTFMLVALYCALAILASKIMTAKENAVDLPFDITFYVYFS 180
Qy 181 LLLIQVLVSCFSDSPFSETIHDNCPCESSAFLSRITFPWITGLIVRGYRPLEGSD 240
Db 181 LLLIQVLVSCFSDSPFSETIHDNCPCESSAFLSRITFPWITGLIVRGYRPLEGSD 240
Qy 241 LMSLNKEDTSQVVPVLYVKWKKECAKTRKQPVKVYSSKQPAQPKSSKYDANEVEAL 300
Db 241 LMSLNKEDTSQVVPVLYVKWKKECAKTRKQPVKVYSSKQPAQPKSSKYDANEVEAL 300
Qy 301 IVKSPQKEMNPSLKVLYKTPGFYFLMSFPFKALHDLMSRGPOLIKLLIFVNDTRAPD 360
Db 301 IVKSPQKEMNPSLKVLYKTPGFYFLMSFPFKALHDLMSRGPOLIKLLIFVNDTRAPD 360
Qy 361 WQGFYTYVLFPVTAQLOTLVLAHQYFHLCEVSGMRKXTAVICAVYRKALVITNSARKSTV 420
Db 361 WQGFYTYVLFPVTAQLOTLVLAHQYFHLCEVSGMRKXTAVICAVYRKALVITNSARKSTV 420
Qy 421 GEIVNLSVDAQRPMDLATYINMTWSAPLQVILALYLLMLNLGPSVLAGVAVMYLMPVN 480
Db 421 GEIVNLSVDAQRPMDLATYINMTWSAPLQVILALYLLMLNLGPSVLAGVAVMYLMPVN 480
Qy 481 AVAMAKTKTYOVAAHKSKDNRIKLMNEILNGIKYLKUYAMELAKDKYLAIOBELKYLK 540
Db 481 AVAMAKTKTYOVAAHKSKDNRIKLMNEILNGIKYLKUYAMELAKDKYLAIOBELKYLK 540
Qy 541 KSAVLSAVGFTWCTPPLVALCTPAVYVTTIDENNIIIDAOTAPSLAFNLRPLNTLP 600
Db 541 KSAVLSAVGFTWCTPPLVALCTPAVYVTTIDENNIIIDAOTAPSLAFNLRPLNTLP 600
Qy 601 MVISIYQASVSLRLRIFLSHEELPDSIRRPVKDGGTNSITVRNATFTWASDPT 660
Db 601 MVISIYQASVSLRLRIFLSHEELPDSIRRPVKDGGTNSITVRNATFTWASDPT 660
Qy 661 LMGITFSPLEGALVAVVGQVCGKSSLLSALLAEMDVKEGHVAIKGSVAVYVPOQAMIOND 720
Db 661 LMGITFSPLEGALVAVVGQVCGKSSLLSALLAEMDVKEGHVAIKGSVAVYVPOQAMIOND 720
Qy 721 SLRENIILFCGLLEBPYRSVYQACALLPDLIELSGDPTLGEKGVNLSGGOKORVSLAR 780
Db 721 SLRENIILFCGLLEBPYRSVYQACALLPDLIELSGDPTLGEKGVNLSGGOKORVSLAR 780

QY 781 AVSNADILYFDDPLSAVDAHVGKHIENNVIGPKGMLKNKTRILLVTHSMSTLPQVDVITV 840
 DB 781 AVSNADILYFDDPLSAVDAHVGKHIENNVIGPKGMLKNKTRILLVTHSMSTLPQVDVITV 840
 QY 841 MSGGKISEMGSYQGLLARDGAPAFPLRTYASTBODAEENGVTGSGPGKEAKOMENGM 900
 DB 841 MSGGKISEMGSYQGLLARDGAPAFPLRTYASTBODAEENGVTGSGPGKEAKOMENGM 900
 QY 901 LVTDSAKOIQOROLSSSSSSSGDISRHHNSTAEIOKEAKEEPMKMEAKAKATQGVKL 960
 DB 901 LVTDSAKOIQOROLSSSSSSSGDISRHHNSTAEIOKEAKEEPMKMEAKAKATQGVKL 960
 QY 961 SVYWDYKKAIGLFIISFLIFLFCNHNVSALASNYLSMTDPIVNGTOEHTKRLSYVG 1020
 DB 961 SVYWDYKKAIGLFIISFLIFLFCNHNVSALASNYLSMTDPIVNGTOEHTKRLSYVG 1020
 QY 1021 ALGISGIIAVFGSMAVSIIGIILASRCLHVDLHSLIRSPMSFFERTPSGNLVNRFSEKL 1080
 DB 1021 ALGISGIIAVFGSMAVSIIGIILASRCLHVDLHSLIRSPMSFFERTPSGNLVNRFSEKL 1080
 QY 1081 DTUDSMIPVYKMGSLFNVTGACIVITLPTAIIIPPLGLIYFFVQAFYVASSROL 1140
 DB 1081 DTUDSMIPVYKMGSLFNVTGACIVITLPTAIIIPPLGLIYFFVQAFYVASSROL 1140
 QY 1141 KRLESVSRSPVYSHFNELLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANWMLA 1200
 DB 1141 KRLESVSRSPVYSHFNELLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANWMLA 1200
 QY 1201 VRLEVCNGLIVLFAALPAVISRHSLSAGVLGVSYSLSQVTTYLNWLVMSSEMETNIVA 1260
 DB 1201 VRLEVCNGLIVLFAALPAVISRHSLSAGVLGVSYSLSQVTTYLNWLVMSSEMETNIVA 1260
 QY 1261 VERIKESSETEKAPMOIQETAPSSWPQVGRVFRNYCYRREDLPVLRHINVTINGG 1320
 DB 1261 VERIKESSETEKAPMOIQETAPSSWPQVGRVFRNYCYRREDLPVLRHINVTINGG 1320
 QY 1321 EKVGIIVGRTGAGKSSLTGLFRINESAGEIINDIGINAKIGLHDLFPKTIITIPQDVLV 1380
 DB 1321 EKVGIIVGRTGAGKSSLTGLFRINESAGEIINDIGINAKIGLHDLFPKTIITIPQDVLV 1380
 QY 1381 SGLRNLMDPFSSQYSDSEVWTSLELAHKDFVSALPKLDHECAEGENLSVGRQVLVCL 1440
 DB 1381 SGLRNLMDPFSSQYSDSEVWTSLELAHKDFVSALPKLDHECAEGENLSVGRQVLVCL 1440
 QY 1441 ARAALRRTKTLVDEAFAAVDLLETDDLIQSTIRIQEFDCVLTIAHRLNTIMDYTRIVL 1500
 DB 1441 ARAALRRTKTLVDEAFAAVDLLETDDLIQSTIRIQEFDCVLTIAHRLNTIMDYTRIVL 1500
 QY 1501 DKGEIOEYGAPSDLLQORGLFYSAKXAGLV 1531
 DB 1501 DKGEIOEYGAPSDLLQORGLFYSAKXAGLV 1531

XX 27-OCT-1992; 92US-00966923.
 PR 08-MAR-1993; 93US-00029340.
 PR 26-OCT-1993; 93US-00141893.
 PR 20-MAR-1995; 95US-00407207.
 XX (TOOH) UNIV QUEBENS KINGSTON.
 PI Cole SPC, Dealey RG;
 XX WPI; 1999-253868/21.
 DR N-PSDB; AAX19818.
 PT Protecting mammalian cells against cytotoxic drugs.
 PS Claim 2; Col 79-86; 82pp; English.
 CC The present sequence represents a human multidrug resistance-associated
 CC protein (MRP). The present invention also describes a method for
 CC protecting a mammalian cell against the cytotoxicity of anthracyclines,
 CC epipodophyllotoxins and Vinca alkaloids (A) by introducing into it a
 CC nucleic acid (I) that hybridizes under stringent conditions to a nucleic
 CC acid (II) that encodes an MRP protein (III). Introduction of (I) protects
 CC cells against cytotoxic effects of (A), particularly to protect normal
 CC cells against (A) being used for treatment of cancers. Cells transformed
 CC with (I) can be used to screen for agents that affect multidrug
 CC resistance or are directly toxic to multidrug resistant cells, i.e.
 CC potential therapeutics for multidrug-resistant cancers. Confering
 CC resistance to normal cells should allow an increase in the dose of (A)
 CC that can be administered safely
 CC XX
 SQ Sequence 1531 AA:
 Query Match 79.4%; Score 7860; DB 2; Length 1531;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MALRGFCSAGSDPLMDMNTNTNTPDFTKCRONTLVVWPCFYLMACPFYLYSRH 60
 DB 1 MALRGFCSAGSDPLMDMNTNTNTPDFTKCRONTLVVWPCFYLMACPFYLYSRH 60
 QY 61 DRGVIQMTPLNKTKTALGFLMTVCMDLTFYSFMRBSRGFLAPVPLVSPTLGITTLA 120
 DB 61 DRGVIQMTPLNKTKTALGFLMTVCMDLTFYSFMRBSRGFLAPVPLVSPTLGITTLA 120
 QY 121 TFLIOERRKGVSSGIMLTFMWALVACALAIIRSKIMTALKEDAQVDFRDIIFYVFS 180
 DB 121 TFLIOERRKGVSSGIMLTFMWALVACALAIIRSKIMTALKEDAQVDFRDIIFYVFS 180
 QY 181 LLLIQVLSGFSRSPFLFSTIHDNPNCPSSASFLSRITFWMITGLIYNGRPLBEGSD 240
 DB 181 LLLIQVLSGFSRSPFLFSTIHDNPNCPSSASFLSRITFWMITGLIYNGRPLBEGSD 240
 QY 241 LMSLNKEDPSEGVVPLVKNWKECAKTRQPKVYSSDPPQPKSSSVVDNNEVEAL 300
 DB 241 LMSLNKEDPSEGVVPLVKNWKECAKTRQPKVYSSDPPQPKSSSVVDNNEVEAL 300
 QY 301 IVKSPQKWNPSLFLKVLTKTFGYPFLMSPFKAIHDLMPSCQIILKLIKFNVDTKAPD 360
 DB 301 IVKSPQKWNPSLFLKVLTKTFGYPFLMSPFKAIHDLMPSCQIILKLIKFNVDTKAPD 360
 QY 361 WQGYFTVLLFVTAQIOTLVLAHQYFHIQVSGMRITAVIGAVYRRLVITNSARKSSTV 420
 DB 361 WQGYFTVLLFVTAQIOTLVLAHQYFHIQVSGMRITAVIGAVYRRLVITNSARKSSTV 420
 QY 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVITATLVMLNTPSVLAGAVVWLVNVPV 480
 DB 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVITATLVMLNTPSVLAGAVVWLVNVPV 480
 QY 481 AVAMKTKTYQVAMHMSKONRIKLMEIILNGIKVLYKLAWELAFAKQKVALAIROBEIKVLK 540
 DB 481 AVAMKTKTYQVAMHMSKONRIKLMEIILNGIKVLYKLAWELAFAKQKVALAIROBEIKVLK 540

```

QY 541 KSAVLSAVGTFTWCTPPLVALCTPAVYVTTIDENNILDAQAFVSLALFNLRLPEPLILP 600
DB 541 KSAVLSAVGTFTWCTPPLVALCTPAVYVTTIDENNILDAQAFVSLALFNLRLPEPLILP 600
QY 601 MVTSSIVQASVSLRLRIPLSHELEPDSIRRRPVKGGCNGNSTVNAATTWARSDDPT 660
DB 601 MVTSSIVQASVSLRLRIPLSHELEPDSIRRRPVKGGCNGNSTVNAATTWARSDDPT 660
QY 661 LMGITFSIPEGALVAVVQVQCGKSSLLSALLAEMDKVEGHVAIKGSAVYPPQQAQOND 720
DB 661 LMGITFSIPEGALVAVVQVQCGKSSLLSALLAEMDKVEGHVAIKGSAVYPPQQAQOND 720
QY 721 SLRENILFGCQLEBPYRSVIQACALLPDLEILPSGDRTEIGKGVNLSGGQKORVSLAR 780
DB 721 SLRENILFGCQLEBPYRSVIQACALLPDLEILPSGDRTEIGKGVNLSGGQKORVSLAR 780
QY 781 AVYENADIYLPDDPLSAVDHVGHIPEENVIGPKGMKNKTRILVTHSMSTLPQVDVIIV 840
DB 781 AVYENADIYLPDDPLSAVDHVGHIPEENVIGPKGMKNKTRILVTHSMSTLPQVDVIIV 840
QY 841 MSGGKISEMGSYQELLARDGAFAEFLRTYASTEOQDAEENGVTGVSQPGKEAKOMENGM 900
DB 841 MSGGKISEMGSYQELLARDGAFAEFLRTYASTEOQDAEENGVTGVSQPGKEAKOMENGM 900
QY 901 LVTDASAGKQIQRQSSSSSSSGDISRHHNSTAELOKAKAKEETWKLMEADKAOTGYKL 960
DB 901 LVTDASAGKQIQRQSSSSSSSGDISRHHNSTAELOKAKAKEETWKLMEADKAOTGYKL 960
QY 961 SVYWDYKAIQGLFTSFLSIFLMCNHVALASNYWLSLWTDPIVNGQHTKRLSYG 1020
DB 961 SVYWDYKAIQGLFTSFLSIFLMCNHVALASNYWLSLWTDPIVNGQHTKRLSYG 1020
QY 1021 ALGISQGIAGVGYMAVSIQGIILASRCLHVDLHLSILSPMSFPERPESGMLVNRFSKEL 1080
DB 1021 ALGISQGIAGVGYMAVSIQGIILASRCLHVDLHLSILSPMSFPERPESGMLVNRFSKEL 1080
QY 1081 DTVDSMIPEVIKMGMSLFENVIGACTIVILLATPIAIIIPGLIYFFVQRPYVASSRQL 1140
DB 1081 DTVDSMIPEVIKMGMSLFENVIGACTIVILLATPIAIIIPGLIYFFVQRPYVASSRQL 1140
QY 1141 KRLESVRSRPYSHFNENETLGVSVTRAFEBQERFTHQDLDVDENOKAYYSIVANRWLA 1200
DB 1141 KRLESVRSRPYSHFNENETLGVSVTRAFEBQERFTHQDLDVDENOKAYYSIVANRWLA 1200
QY 1201 VRLECVGNCIYLPALFAVISRHSISAGLVGSYSIQVTTYLNMVLRMSSEMETNIVA 1260
DB 1201 VRLECVGNCIYLPALFAVISRHSISAGLVGSYSIQVTTYLNMVLRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPWQIOETAPPSWPQVRVEFRNYCLARYREDLDFVLRIHNTVINGG 1320
DB 1261 VERLKEYSETEKAPWQIOETAPPSWPQVRVEFRNYCLARYREDLDFVLRIHNTVINGG 1320
QY 1321 EKVGIVGRTGAGKSLTIGLFRINESAGEIITIDGINIAKIGLHDLRKITITIPQDPYLF 1380
DB 1321 EKVGIVGRTGAGKSLTIGLFRINESAGEIITIDGINIAKIGLHDLRKITITIPQDPYLF 1380
QY 1381 SSGSRMNLDPSSQSDSEEWMTSLBLAHKDFVSALPDLDHBCAGGENTLVSQORQLVCL 1440
DB 1381 SSGSRMNLDPSSQSDSEEWMTSLBLAHKDFVSALPDLDHBCAGGENTLVSQORQLVCL 1440
QY 1441 ARALLRKTKIILVDEATAVAVDLETDLLIQSTIRIQFEDCTVLTIAHRLNTIMDYTRIVL 1500
DB 1441 ARALLRKTKIILVDEATAVAVDLETDLLIQSTIRIQFEDCTVLTIAHRLNTIMDYTRIVL 1500
QY 1501 DKGEIOEGARSDLIQORGLFYSMAXKDGVL 1531
DB 1501 DKGEIOEGARSDLIQORGLFYSMAXKDGVL 1531

```

RESULT 8
AA55799
ID AA55799 standard; protein; 1531 AA.
XX

```

AC AA55799;
XX 28-FEB-2000 (first entry)
DT XX
DE Human multidrug resistance-associated protein (MRP) variant.
XX
KW Chemosensitizer; multidrug resistance-associated protein; MRP; human;
KW therapeutic agent; P-glycoprotein-mediated multidrug resistance; lung;
KW cancer; variant.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 685
FT /label= L685S
FT /note= "wild-type Leu is replaced by Ser"
FT Misc-difference 1282
FT /label= R1282A
FT /note= "wild-type Arg is replaced by Ala"
XX
PN US6001563-A.
XX
PD 14-DEC-1999.
XX
PE 05-JUN-1995; 95US-00463179.
XX
PR 27-OCT-1992; 92US-00966923.
PR 08-MAR-1993; 93US-00029340.
PR 26-OCT-1993; 93US-00141893.
PR 20-MAR-1995; 95US-00407207.
XX
PA (TOOH ) UNIV QUEBENS KINGSTON.
XX
PI Cole SP, Deeley RG;
XX
DR WPI; 2000-061877/05.
DR N-PSDB; AA339556.
XX
PT Identification of chemosensitizers useful for treating cancer, using
PT nucleic acid encoding multidrug resistance-associated protein.
XX
PS Claim 3; Col 69-80; 77pp; English.
XX
CC The invention provides a method for identifying a substance which is a
CC chemosensitizer that comprises, contacting a cell transfected with
CC nucleic acid encoding multidrug resistance-associated protein (MRP) with
CC a therapeutic agent in vitro. The method is useful for identifying
CC chemosensitizers which may then be used to treat cancer (especially lung
CC cancer). The method allows the identification of chemosensitizers which
CC do not reverse P-glycoprotein-mediated multidrug resistance. The present
CC sequence represents a human MRP variant
XX
SQ Sequence 1531 AA.

```

Query Match 79.4%; Score 7860; DB 3; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MALMGFCSADSDPLMDMNTVMTNSNDPTKCFQNTYLVWPCGYLWACFPFYLYLSRH 60
DB 1 MALMGFCSADSDPLMDMNTVMTNSNDPTKCFQNTYLVWPCGYLWACFPFYLYLSRH 60
QY 61 DRGYIOMTPLNKKTALGFLIMVCMADLFYSPERSRGIFLAVFLVSPFLIGITTLA 120
DB 61 DRGYIOMTPLNKKTALGFLIMVCMADLFYSPERSRGIFLAVFLVSPFLIGITTLA 120
QY 121 TFLIQLERRKGVQSSGIMLTFWLVAVCALAIRSKIMTALKEDAQVDLFFDITFYVYFS 180
DB 121 TFLIQLERRKGVQSSGIMLTFWLVAVCALAIRSKIMTALKEDAQVDLFFDITFYVYFS 180
QY 181 LLLIQVLVSCFSDSPFLFSETIHDNPPCPSSASFLSRITTWITGULVRGYRQPLESSD 240
DB 181 LLLIQVLVSCFSDSPFLFSETIHDNPPCPSSASFLSRITTWITGULVRGYRQPLESSD 240

```

QY 241 LMSLNKEDTSEOVVPLVVKWKKCECAKTRKOPVKVYSSKDPAPPKSSKVDANEVEAL 300
 DB 241 LMSLNKEDTSEOVVPLVVKWKKCECAKTRKOPVKVYSSKDPAPPKSSKVDANEVEAL 300
 QY 301 IVSPQKEMNPSLFKVLKFTGFPYFLMSFFPKAIDLMFSGPOLIKLLIKFVNDTRAPD 360
 DB 301 IVSPQKEMNPSLFKVLKFTGFPYFLMSFFPKAIDLMFSGPOLIKLLIKFVNDTRAPD 360
 QY 361 WQGFYVVLFFVACLOTLVLHOFHICFVSGMKIKTAVIGAVYRKALVTINSARKSSTV 420
 DB 361 WQGFYVVLFFVACLOTLVLHOFHICFVSGMKIKTAVIGAVYRKALVTINSARKSSTV 420
 QY 421 GEIYNLMSVDAQRMDLATYINMIMSAPIQVTLALYLLMLNGPSVLAVAVMLAMPVN 480
 DB 421 GEIYNLMSVDAQRMDLATYINMIMSAPIQVTLALYLLMLNGPSVLAVAVMLAMPVN 480
 QY 481 AVNMMKTXYQVAMHMSKDNRIKLMNEILNGIKVLYAMELAEKDKVLAIROBELKVLK 540
 DB 481 AVNMMKTXYQVAMHMSKDNRIKLMNEILNGIKVLYAMELAEKDKVLAIROBELKVLK 540
 QY 541 KSAVLAVGTFTWCTPPLVALCTPRAVYITIDENNIIDAQTAFLVSLFNLIRPILNP 600
 DB 541 KSAVLAVGTFTWCTPPLVALCTPRAVYITIDENNIIDAQTAFLVSLFNLIRPILNP 600
 QY 601 MVISSIVQASVSLKRLIFLSHELEBDSIERRPVKDGGGNSITVNNATFTARSDPT 660
 DB 601 MVISSIVQASVSLKRLIFLSHELEBDSIERRPVKDGGGNSITVNNATFTARSDPT 660
 QY 661 LNCITFESIPGALVAVVGVCCKSSLSALLAEMDKYEGHVAIKGSVAVYPOQAWQND 720
 DB 661 LNCITFESIPGALVAVVGVCCKSSLSALLAEMDKYEGHVAIKGSVAVYPOQAWQND 720
 QY 721 SLRENIILFGQLEBPYRSYIQAQALLPDLIELIPSGDRTBEIGEXVNLSEGQKQVSLAR 780
 DB 721 SLRENIILFGQLEBPYRSYIQAQALLPDLIELIPSGDRTBEIGEXVNLSEGQKQVSLAR 780
 QY 781 AAVSNMDIYLPDPLASVDAHVGHIFENYIGPKMCKNTRILLVHSMYSLPOVDYIV 840
 DB 781 AAVSNMDIYLPDPLASVDAHVGHIFENYIGPKMCKNTRILLVHSMYSLPOVDYIV 840
 QY 841 MSGGKISEMGSYOELARDGAFAEFLRTYASTBOEODAEENGVTGVSQPKGAQOMENG 900
 DB 841 MSGGKISEMGSYOELARDGAFAEFLRTYASTBOEODAEENGVTGVSQPKGAQOMENG 900
 QY 901 LVTDGAKQLOQOLSSSSSYSGDISRHNSSTAELQKAEAKKEETWKLMEADKAQTQVKL 960
 DB 901 LVTDGAKQLOQOLSSSSSYSGDISRHNSSTAELQKAEAKKEETWKLMEADKAQTQVKL 960
 QY 961 SYVWDYKATIGLFTSPLSTFLPMONHVSALASNYMLSLMTDDPIVNGTOHTKRLSVYG 1020
 DB 961 SYVWDYKATIGLFTSPLSTFLPMONHVSALASNYMLSLMTDDPIVNGTOHTKRLSVYG 1020
 QY 1021 ALGISOGIAVFGYSMAVISGIIASRCLAHVDLHSIIRSPMSFPERTPSGNLVNRSKEL 1080
 DB 1021 ALGISOGIAVFGYSMAVISGIIASRCLAHVDLHSIIRSPMSFPERTPSGNLVNRSKEL 1080
 QY 1081 DTVDSMIPVYIKMPMGLFNVIGACIVILLATPIAIIIPPLGLIYFPVORFYVASSROL 1140
 DB 1081 DTVDSMIPVYIKMPMGLFNVIGACIVILLATPIAIIIPPLGLIYFPVORFYVASSROL 1140
 QY 1141 KRLAESYRSFVYGHFNFTLLGVSVIRAFPEOERFIHOSDKVDENQAAVYPSIVANMWA 1200
 DB 1141 KRLAESYRSFVYGHFNFTLLGVSVIRAFPEOERFIHOSDKVDENQAAVYPSIVANMWA 1200
 QY 1201 VRLCEVNCICVLPAALFAVISRHSLSAGVLGSYSLSQVTTYLMNLVRSSEMETIVA 1260
 DB 1201 VRLCEVNCICVLPAALFAVISRHSLSAGVLGSYSLSQVTTYLMNLVRSSEMETIVA 1260
 QY 1261 VERIKESYSEKAPWQIOETAPSSWPQVGRVFNRYCLIRREDDLPVLRHINVTINGG 1320
 DB 1261 VERIKESYSEKAPWQIOETAPSSWPQVGRVFNRYCLIRREDDLPVLRHINVTINGG 1320
 QY 1321 EKVIGVGRGAKSSLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
 DB 1321 EKVIGVGRGAKSSLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
 QY 1381 SSGILRMNLDPFSQYSDSEVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGRQOLVCL 1440
 DB 1381 SSGILRMNLDPFSQYSDSEVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGRQOLVCL 1440
 QY 1441 ARALLRRTKTLVDEATAAADLETDDLIQSTIFRQFEDCTVLTAHRLNTIMOTRYIVL 1500
 DB 1441 ARALLRRTKTLVDEATAAADLETDDLIQSTIFRQFEDCTVLTAHRLNTIMOTRYIVL 1500
 QY 1501 DKGEIOEYGAAPSDILOQRLGYFSMAKADAGLV 1531
 DB 1501 DKGEIOEYGAAPSDILOQRLGYFSMAKADAGLV 1531

QY 1321 EKVIGVGRGAKSSLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
 DB 1321 EKVIGVGRGAKSSLTGLFRINESAGEBIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
 QY 1381 SSGILRMNLDPFSQYSDSEVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGRQOLVCL 1440
 DB 1381 SSGILRMNLDPFSQYSDSEVWTSLELAHLKDFVSALPDKLDHECAEGENLSVGRQOLVCL 1440
 QY 1441 ARALLRRTKTLVDEATAAADLETDDLIQSTIFRQFEDCTVLTAHRLNTIMOTRYIVL 1500
 DB 1441 ARALLRRTKTLVDEATAAADLETDDLIQSTIFRQFEDCTVLTAHRLNTIMOTRYIVL 1500
 QY 1501 DKGEIOEYGAAPSDILOQRLGYFSMAKADAGLV 1531
 DB 1501 DKGEIOEYGAAPSDILOQRLGYFSMAKADAGLV 1531

RESULT 9
 AA78873
 ID AA78873 standard; protein; 1531 AA.
 XX
 AC AA78873;
 XX
 DT 19-MAY-2000 (first entry)
 XX
 DE Multidrug resistance protein (MRP) natural variant amino acid sequence.
 XX
 KM Multidrug resistance protein; MRP; human; anthracycline; Vinca alkaloid;
 KW epipodophyllotoxin; cancer; leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN US6025473-A.
 XX
 PD 15-FEB-2000.
 XX
 PF 05-JUN-1995; 95US-00461384.
 XX
 PR 27-OCT-1992; 92US-00966923.
 PR 08-MAR-1993; 93US-00029340.
 PR 26-OCT-1993; 93US-00141893.
 PR 20-MAR-1995; 95US-00407207.
 XX
 PA (TOOH) UNIV QUEBENS KINGSTON.
 XX
 PI Cole SPC, Deeley RG;
 XX
 DR WPI; 2000-181838/16.
 DR N-PDB; AA290193.
 XX
 PT Isolated protein conferring multidrug resistance, to at least two drugs
 PT selected from anthracyclines, epipodophyllotoxins and Vinca alkaloids, on
 PT a drug sensitive mammalian cell.
 XX
 PS Claim 10; Col 79-88; 78pp; English.
 XX
 CC This sequence represents a human multidrug resistance protein (MRP)
 CC natural variant amino acid sequence. The human MRP confers multidrug
 CC resistance, including resistance to at least two drugs selected from
 CC anthracyclines, epipodophyllotoxins and Vinca alkaloids, on a drug
 CC sensitive mammalian cell, when the protein is expressed in the cell. The
 CC multidrug resistance is not substantially reversed by chemosensitizers
 CC which reverse P-glycoprotein-mediated multidrug resistance. The MRP
 CC protein sequence can be used to generate antibodies against MRP. The MRP
 CC protein and nucleotide sequences can be used in compositions which are
 CC used to treat patients with tumours displaying multidrug resistance. The
 CC compositions and methods of the invention can be used particularly to
 CC treat breast cancer, leukaemia, fibrosarcomas, cervical cancer, gliomas,
 CC thymoma, neuroblastoma, and lung cancer. Antibodies directed against
 CC MRP can be used to inhibit the multidrug resistance of a multidrug
 CC resistant cell
 XX
 SQ Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 3; Length 1531;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MALRGFSSAGSDLDLMDNMNTWNTSNPDFTCFQNTVLVWPCFYLMACPFYLYLSRH 60
DB 1 MALRGFSSAGSDLDLMDNMNTWNTSNPDFTCFQNTVLVWPCFYLMACPFYLYLSRH 60
QY 61 DRGIQNTPLINKTALGFLLMIVCMADLFYSFMRSGIFLAPVFLVSPILLGITLLA 120
DB 61 DRGIQNTPLINKTALGFLLMIVCMADLFYSFMRSGIFLAPVFLVSPILLGITLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFWLVALYCALAIILRSKIMTALKEDAQVDLDFDITFYVYS 180
DB 121 TFLIQLERRKGVSSGIMLTFWLVALYCALAIILRSKIMTALKEDAQVDLDFDITFYVYS 180
QY 181 LLLIQLVLSGSDSPLEFSETHDNPCESSASFLSITFWMTGLIVRGYRQPLBESD 240
DB 181 LLLIQLVLSGSDSPLEFSETHDNPCESSASFLSITFWMTGLIVRGYRQPLBESD 240
QY 241 LMSLNKEDTSEQVVPVLVKMKKCECAKTRKQPVKVVYSSKDPAPKESKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKMKKCECAKTRKQPVKVVYSSKDPAPKESKVDANEVEAL 300
QY 301 IVKSPKEMWNSLFKVLKTEGYPFLMSFPFKAJHDLMMFSGPOLKLLIKFVNDTRAPD 360
DB 301 IVKSPKEMWNSLFKVLKTEGYPFLMSFPFKAJHDLMMFSGPOLKLLIKFVNDTRAPD 360
QY 361 WQGFYVVLVFTVACLQTLVHOYFHICFVSGMRKKTAVIGAVRKALVITNSARKSTV 420
DB 361 WQGFYVVLVFTVACLQTLVHOYFHICFVSGMRKKTAVIGAVRKALVITNSARKSTV 420
QY 421 GEIYNLMSVDQRFPMDLATYINMIWSAPLOYILALYLLMLNGSPVLGAVAVMLVNVN 480
DB 421 GEIYNLMSVDQRFPMDLATYINMIWSAPLOYILALYLLMLNGSPVLGAVAVMLVNVN 480
QY 481 AVMAAMKTQYVAAHMKSKDNRIKLMNEILLNGIKVLKLYAMEIAFDKYLAIROBELKYLK 540
DB 481 AVMAAMKTQYVAAHMKSKDNRIKLMNEILLNGIKVLKLYAMEIAFDKYLAIROBELKYLK 540
QY 541 KSAVLASAVGTFTWCTPFLVALCTPAVYVITDENNILLDAQNAFSLALFNILRPPLNLP 600
DB 541 KSAVLASAVGTFTWCTPFLVALCTPAVYVITDENNILLDAQNAFSLALFNILRPPLNLP 600
QY 601 MVISIYOASVSLRLRIFLSHELEPDSIERRPVKGSGNSTVRATFTMARSDPT 660
DB 601 MVISIYOASVSLRLRIFLSHELEPDSIERRPVKGSGNSTVRATFTMARSDPT 660
QY 661 LMGITFSIPGALVAVVGVGCGKSSLLSALLAEMDKYEGHVAIKGSVAVYVPOQAMIOND 720
DB 661 LMGITFSIPGALVAVVGVGCGKSSLLSALLAEMDKYEGHVAIKGSVAVYVPOQAMIOND 720
QY 721 SLRENIILFGCOLLEBPYRSVIOACALPLDLEILPSGDRTEIGEKVNLSGGOKORVSLAR 780
DB 721 SLRENIILFGCOLLEBPYRSVIOACALPLDLEILPSGDRTEIGEKVNLSGGOKORVSLAR 780
QY 781 AVYNNADLYLPDDPLSAVDAHVGKHIFENVIQPKMLNKKRIILVTHSMYLPQVDVITV 840
DB 781 AVYNNADLYLPDDPLSAVDAHVGKHIFENVIQPKMLNKKRIILVTHSMYLPQVDVITV 840
QY 841 MSGGKISEMSGYOELIARDGAPAEFLRYASTEQODAEENGVTGVSFGKEAKOMENGM 900
DB 841 MSGGKISEMSGYOELIARDGAPAEFLRYASTEQODAEENGVTGVSFGKEAKOMENGM 900
QY 901 LVTDASAGKOLQROUSSSSSYSGDISRHNSSTAELOKAPAKKEETWKLMEADQAQGYKL 960
DB 901 LVTDASAGKOLQROUSSSSSYSGDISRHNSSTAELOKAPAKKEETWKLMEADQAQGYKL 960
QY 961 SVYDYMKAIGLFTSFLSIFLPMCHVSLASNYWLSLMTDPIVNGQEHRTKVLSTYNG 1020
DB 961 SVYDYMKAIGLFTSFLSIFLPMCHVSLASNYWLSLMTDPIVNGQEHRTKVLSTYNG 1020

```

```

QY 1021 ALGISGIAVFGYSMAVSGGIIASRCLHYDLLHSIIRSMSPFERPTPSGNLVNREKEL 1080
DB 1021 ALGISGIAVFGYSMAVSGGIIASRCLHYDLLHSIIRSMSPFERPTPSGNLVNREKEL 1080
QY 1081 DTDSMTPEVIKFMGSLFENVIGACTVILATPAIIIPPLIGLYFVVOFPYASSROL 1140
DB 1081 DTDSMTPEVIKFMGSLFENVIGACTVILATPAIIIPPLIGLYFVVOFPYASSROL 1140
QY 1141 KRLESVRSFVYSHFNELLGVSIVYRAFEQERFIHOSDKVDENQAYYPSIVANRMLA 1200
DB 1141 KRLESVRSFVYSHFNELLGVSIVYRAFEQERFIHOSDKVDENQAYYPSIVANRMLA 1200
QY 1201 VRLFCVNCIVLPALPAVYSRHSLSAGLVGSYSLOVTTYINLVYRMSSEMETNIVA 1260
DB 1201 VRLFCVNCIVLPALPAVYSRHSLSAGLVGSYSLOVTTYINLVYRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPPSSMPQVGRVFRNYCLRYREDLPFVLRHIVTINGG 1320
DB 1261 VERLKEYSETEKEAPWQIOETAPPSSMPQVGRVFRNYCLRYREDLPFVLRHIVTINGG 1320
QY 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVL 1380
DB 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVL 1380
QY 1381 SSGIRANLDPSSQYSDDEEVTSLFLAHLKQFVSALPKLHCEKAGEENLSVGOROLVCL 1440
DB 1381 SSGIRANLDPSSQYSDDEEVTSLFLAHLKQFVSALPKLHCEKAGEENLSVGOROLVCL 1440
QY 1441 ABALRRTKTLIVUDEAFAVADLEETDDLIOSTIRFOPEDCVLTIAHRLNTIMDTRIVL 1500
DB 1441 ABALRRTKTLIVUDEAFAVADLEETDDLIOSTIRFOPEDCVLTIAHRLNTIMDTRIVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531

```

```

RESULT 10
ABG61810
ID ABG61810 standard; protein; 1531 AA.
XX
AC ABG61810;
XX
DT 15-AUG-2002 (first entry)
XX
DE Prostate cancer-associated protein #11.
XX
KM Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
XX
OS Mammalia.
XX
PN M0200230268-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US032045.
XX
PR 13-OCT-2000; 2000US-00687576.
PR 08-DEC-2000; 2000US-00733388.
PR 08-DEC-2000; 2000US-00733742.
PR 24-JAN-2001; 2001US-0263957P.
PR 16-MAR-2001; 2001US-0276791P.
PR 16-MAR-2001; 2001US-0276888P.
PR 06-APR-2001; 2001US-0281922P.
PR 24-APR-2001; 2001US-0286214P.
PR 30-APR-2001; 2001US-00847046.
PR 04-MAY-2001; 2001US-0288589P.
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX
WP1; 2002-471335/50.

```

DR N-PSDB; ABK92125.
XX Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.
PS
XX Claim 27; Page 309; 436pp; English.
XX
CC The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC associating a biological sample from the patient with prostate cancer-
CC associated polynucleotides (designated PC genes) that selectively
CC hybridize to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins
CC
XX
SQ Sequence 1531 AA;
Query Match 79.4%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFSGADSDPLDMNMTNTNPNPTFCQNTVLYWVPCRYLWACPFYFLYSRH 60
DB 1 MALRGFSGADSDPLDMNMTNTNPNPTFCQNTVLYWVPCRYLWACPFYFLYSRH 60
QY 61 DRGVIQMTPLNKITAGFLIMIVCMADLFYSFMERSRGIFLAFVFLVPTLGIITLLA 120
DB 61 DRGVIQMTPLNKITAGFLIMIVCMADLFYSFMERSRGIFLAFVFLVPTLGIITLLA 120
QY 121 TPLIQLERRRGVSSGIMLTFWLVVALCALILRSKIMTALKEDAQVDLFFDITFYVFS 180
DB 121 TPLIQLERRRGVSSGIMLTFWLVVALCALILRSKIMTALKEDAQVDLFFDITFYVFS 180
QY 181 LLLIQLVLSGCSDDSPLESETIHDNPNCPRESSASFLSRITFWMTTGLIVRGYRPLGSD 240
DB 181 LLLIQLVLSGCSDDSPLESETIHDNPNCPRESSASFLSRITFWMTTGLIVRGYRPLGSD 240
QY 241 LMSLNKEDTSEQVVPVLYKWKKECAKTRKQPVYVYSSKDPAPKSSKVDANEYEAL 300
DB 241 LMSLNKEDTSEQVVPVLYKWKKECAKTRKQPVYVYSSKDPAPKSSKVDANEYEAL 300
QY 301 IVKSPQKEMNPSEFKVLYKTGPFYLMSPFKALHDLMPGSGPOLIKFVNDTRAPD 360
DB 301 IVKSPQKEMNPSEFKVLYKTGPFYLMSPFKALHDLMPGSGPOLIKFVNDTRAPD 360
QY 361 WQGFYVVLFPVACIOTLVLAHQVYHICFVSGMRTKAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYVVLFPVACIOTLVLAHQVYHICFVSGMRTKAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIYNLMSVDAQRFMDLATYINMISAPLOVITLALYLLMLNGSGVLAGVAVMLAMPVN 480
DB 421 GEIYNLMSVDAQRFMDLATYINMISAPLOVITLALYLLMLNGSGVLAGVAVMLAMPVN 480
QY 481 AVNMAKTKTYQVAMHMSKDNRIKIMNEILNGIKVLYLAMELAFKDYALIRQELKYLK 540
DB 481 AVNMAKTKTYQVAMHMSKDNRIKIMNEILNGIKVLYLAMELAFKDYALIRQELKYLK 540
QY 541 KSAVLSAVGTFTWCTPEIVLACFEAVYVITDENNNIDAGTAFVSLFENILRPPLNLP 600
DB 541 KSAVLSAVGTFTWCTPEIVLACFEAVYVITDENNNIDAGTAFVSLFENILRPPLNLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEELPDSIERRPVKDGGGNSITVNNATFTMARSDPPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEELPDSIERRPVKDGGGNSITVNNATFTMARSDPPT 660

DB 601 MVISSIVQASVSLKRLRIFLSHEELPDSIERRPVKDGGGNSITVNNATFTMARSDPPT 660
QY 661 LMGITTSIPREGALVAVYGVGCCGSSLLSALLAMDMYVEGHVAKGSVAVYPOQAWIQND 720
DB 661 LMGITTSIPREGALVAVYGVGCCGSSLLSALLAMDMYVEGHVAKGSVAVYPOQAWIQND 720
QY 721 SLRENIIFGCOLSEPPYRSVYIQAACALPDLEILPDSGRTEIGEGVNLSSGQKQVSLAR 780
DB 721 SLRENIIFGCOLSEPPYRSVYIQAACALPDLEILPDSGRTEIGEGVNLSSGQKQVSLAR 780
QY 781 AVYSNADIVYFDDPLSADVDAHVGKHIPEENVYIGPKMKLKNKTRILVTHSMSTLPQVDYIIV 840
DB 781 AVYSNADIVYFDDPLSADVDAHVGKHIPEENVYIGPKMKLKNKTRILVTHSMSTLPQVDYIIV 840
QY 841 MSGGKIEMSGSYQCLLARDGAFAPFLATYASTEBQDAEENGVTGSGPGKEAQMENG 900
DB 841 MSGGKIEMSGSYQCLLARDGAFAPFLATYASTEBQDAEENGVTGSGPGKEAQMENG 900
QY 901 LVTDASQKOLQORLSSSSSYSGDISRHNSTAELQKAKKEETWKLMEADKATQGVKL 960
DB 901 LVTDASQKOLQORLSSSSSYSGDISRHNSTAELQKAKKEETWKLMEADKATQGVKL 960
QY 961 SYVWDYMKALGLFISFLSIFLFCMCHVSALASNYMSTLMTDDPIVNGTOEHTKRLSVYG 1020
DB 961 SYVWDYMKALGLFISFLSIFLFCMCHVSALASNYMSTLMTDDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISQIAVFGYSMAVSIQIILASRCLAYDLHSILRSPMSFEPTPSGNLVNRFSEKL 1080
DB 1021 ALGISQIAVFGYSMAVSIQIILASRCLAYDLHSILRSPMSFEPTPSGNLVNRFSEKL 1080
QY 1081 DTWDSMIPEYIKMFMGSLFNVGACIVITLATAPIIIPPLGITYFVORFYASSROL 1140
DB 1081 DTWDSMIPEYIKMFMGSLFNVGACIVITLATAPIIIPPLGITYFVORFYASSROL 1140
QY 1141 KRLSEYSRSPVYSHFNFTLLGVSIVIRAFEEQERFIHOSDLKVDENQKAYPSIVANMFLA 1200
DB 1141 KRLSEYSRSPVYSHFNFTLLGVSIVIRAFEEQERFIHOSDLKVDENQKAYPSIVANMFLA 1200
QY 1201 VRLCEVNCIVLPALFAVISRHSLSAGLVGSVSYSLQVTTYLNLVMSSEMETNIVA 1260
DB 1201 VRLCEVNCIVLPALFAVISRHSLSAGLVGSVSYSLQVTTYLNLVMSSEMETNIVA 1260
QY 1261 VERLKEYSERTEKAPMIOETAPSSWPQVGRVEFRVYCLRYREDLDPVLRHINVTING 1320
DB 1261 VERLKEYSERTEKAPMIOETAPSSWPQVGRVEFRVYCLRYREDLDPVLRHINVTING 1320
QY 1321 EKYGIVGRTGAKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKYGIVGRTGAKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SSGILRNMLDPPSOYSDEEVTSLLELAHLKDFVSALPDKLDHECAEGENLSVQORQVCL 1440
DB 1381 SSGILRNMLDPPSOYSDEEVTSLLELAHLKDFVSALPDKLDHECAEGENLSVQORQVCL 1440
QY 1441 ARALLRKTKLVLDEATAAVALDETDDLIOGTITIQEEDCTVLTARHLANTIMYTRIVL 1500
DB 1441 ARALLRKTKLVLDEATAAVALDETDDLIOGTITIQEEDCTVLTARHLANTIMYTRIVL 1500
QY 1501 DKGEIOEYGAPSDLOOQGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGAPSDLOOQGLFYSMAKDAGLV 1531
RESULT 11
ABM35012
ID ABM35012 standard; protein; 1531 AA.
XX
AC ABM35012;
XX
DT 08-OCT-2003 (first entry)
XX
DE Cancer based on CYP3A5 related protein SEQ ID NO:678.
XX

KM Cancer; CYP3A5; irinotecan; pharmaceutical; malignant glioma;
KM cytochrome p450; subfamily IIIA; nifedipine oxidase; polypeptide 5;
KM cytostratic.
XX Unidentified.
XX
XX WO2003013534-A2.
XX
XX 20-FEB-2003.
XX
XX 23-JUL-2002; 2002WO-EP008219.
XX
XX 23-JUL-2001; 2001EP-00117608.
XX
XX 24-MAY-2002; 2002EP-00011710.
XX
XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
XX Heinrich G, Kerb R;
XX
XX WPI; 2003-268144/26.
XX
XX
XX New use of irinotecan for preparation of compositions for treating cancer
PT in subject having genome with variant allele comprising cytochrome p450,
PT subfamily IIIA, polypeptide 5 polynucleotide, termed CYP3A5.
XX
XX Disclosure; SEQ ID NO 678; 86pp; English.
XX
XX
XX The present invention describes the use of irinotecan (I) or its
CC derivative for the preparation of a pharmaceutical composition for
CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic
CC cancer, or malignant glioma in a subject having a genome with a variant
CC allele comprising a cytochrome p450, subfamily IIIA (nifedipine
CC oxidase), polypeptide 5 (CYP3A5) polynucleotide (II). (I) and (II) have
CC cytostratic activity. The therapeutic applications of (I) is improved,
CC since it is possible to individually treat a subject with an appropriate
CC dosage and/or an appropriate derivative of (I). Therefore, undesirable,
CC harmful or toxic effects are efficiently avoided. Unnecessary and
CC potentially harmful treatment of those subjects who do not respond to the
CC treatment with substances (nonresponders), as well as the development of
CC drug resistances due to suboptimal drug dosing can be avoided. ACF62200
CC to ACF62751 and ABM34912 to ABM35013 represent sequences used in the
CC exemplification of the present invention
XX
XX
XX Sequence 1531 AA;
SQ
Query Match 79.4%; Score 7860; DB 6; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALNGFCSDGSDPLMDNMTNTNSPDTKCFONTVLWVPCFYLMACFPPEFYLYLSRH 60
DB 1 MALNGFCSDGSDPLMDNMTNTNSPDTKCFONTVLWVPCFYLMACFPPEFYLYLSRH 60
QY 61 DRGVIQMTPLNKTALGFLIMIVCWADLFYSFWRSGIFLAPVFLVSPITLGLITTLA 120
DB 61 DRGVIQMTPLNKTALGFLIMIVCWADLFYSFWRSGIFLAPVFLVSPITLGLITTLA 120
QY 121 TFLQLERKRGVQSSGIMLTFPLVALVCALALRSKIMTALKEQOVDFRITTYVYS 180
DB 121 TFLQLERKRGVQSSGIMLTFPLVALVCALALRSKIMTALKEQOVDFRITTYVYS 180
QY 181 LLLLIQVLVSCFSDRPLFSETIHDNPNCPRESSASFLSRITTFWMIIGLIVRGYRQLESD 240
DB 181 LLLLIQVLVSCFSDRPLFSETIHDNPNCPRESSASFLSRITTFWMIIGLIVRGYRQLESD 240
QY 241 LMSLNKEDTSSQVVPVLVKNMKKECAKTRKQPVKVVYSSKDBAQPKSSKVDANEVVAL 300
DB 241 LMSLNKEDTSSQVVPVLVKNMKKECAKTRKQPVKVVYSSKDBAQPKSSKVDANEVVAL 300
QY 301 IVKSGQKEMNBSLFLVLYKTGPFYFLMSFFPKAIIHDLMMFSGPQIIKLKIFVNDTKAPD 360
DB 301 IVKSGQKEMNBSLFLVLYKTGPFYFLMSFFPKAIIHDLMMFSGPQIIKLKIFVNDTKAPD 360

QY 361 WQGYFTVLLFVTAQTLVLAHOYFHIICVSGMKIKTAIVGAVRKALVITNSARKSSTV 420
DB 361 WQGYFTVLLFVTAQTLVLAHOYFHIICVSGMKIKTAIVGAVRKALVITNSARKSSTV 420
QY 421 GEIVNLSVDAQRMIDATYINMTWSAPLOVITLALYLMNIGPSVLAGVAVWVLAMPVN 480
DB 421 GEIVNLSVDAQRMIDATYINMTWSAPLOVITLALYLMNIGPSVLAGVAVWVLAMPVN 480
QY 481 AVNMAKTKTYQVAMKSKDNRIKLMNEILINGIKVILKLYAMELAFKDYVALROBELKYLK 540
DB 481 AVNMAKTKTYQVAMKSKDNRIKLMNEILINGIKVILKLYAMELAFKDYVALROBELKYLK 540
QY 541 KSAVLSAVGFTTWCCTPPLVALCTPAVYVTIIDENNIIIDAOTAFVSLAFNLRPLNLP 600
DB 541 KSAVLSAVGFTTWCCTPPLVALCTPAVYVTIIDENNIIIDAOTAFVSLAFNLRPLNLP 600
QY 601 MVISIIQASVSLRLRIPLSHELEPDSIERREVKGGGNSITVRNAITFMARSDEPT 660
DB 601 MVISIIQASVSLRLRIPLSHELEPDSIERREVKGGGNSITVRNAITFMARSDEPT 660
QY 661 LNCITFSIPREGALVAVVGVCGCKSLSLALLAEMDKVEGHVAIKGSVAVYPOQAWIOND 720
DB 661 LNCITFSIPREGALVAVVGVCGCKSLSLALLAEMDKVEGHVAIKGSVAVYPOQAWIOND 720
QY 721 SLRENIIIFGCOLLEPPYRSVIQACALLPDLEILPSGRTBI GEKGVNLSGQKQVSLAR 780
DB 721 SLRENIIIFGCOLLEPPYRSVIQACALLPDLEILPSGRTBI GEKGVNLSGQKQVSLAR 780
QY 781 AVYSNADIYLFDDLSAVDAHVGHIFENYIGPMKNKTRILVTHSMSTLPVDVYIV 840
DB 781 AVYSNADIYLFDDLSAVDAHVGHIFENYIGPMKNKTRILVTHSMSTLPVDVYIV 840
QY 841 MSGGKISEMGSYQELLARDGAFAEFLRTYASTEOBQDAEENGVTGVSQPKAEAKOMENG 900
DB 841 MSGGKISEMGSYQELLARDGAFAEFLRTYASTEOBQDAEENGVTGVSQPKAEAKOMENG 900
QY 901 LVYDSAGKOLQROUSSSSYSYSGDISRHNSYABLOKAEKKEFTWKIMEADKAOYGVL 960
DB 901 LVYDSAGKOLQROUSSSSYSYSGDISRHNSYABLOKAEKKEFTWKIMEADKAOYGVL 960
QY 961 SYVMDYKATGLFISFISIFLFCMNHVSALASNYWLSLMTDDPLVNGQOHTKRLSYG 1020
DB 961 SYVMDYKATGLFISFISIFLFCMNHVSALASNYWLSLMTDDPLVNGQOHTKRLSYG 1020
QY 1021 ALGISOGIAVFGYMAVSIIGIILASRCHVDLHLSILSPMSFEPTPSGNLVNRFSEL 1080
DB 1021 ALGISOGIAVFGYMAVSIIGIILASRCHVDLHLSILSPMSFEPTPSGNLVNRFSEL 1080
QY 1081 DTVDSMIPEVIKMGMSLFNVIGACIYILLATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
DB 1081 DTVDSMIPEVIKMGMSLFNVIGACIYILLATPIAIIIPPLGLIYFFVQRFYVASSROL 1140
QY 1141 KRLESVSRSPPYSHFNETLIGVSVIARAEDEBERIHOSSDLKVDENQAXYYSIVANRPLA 1200
DB 1141 KRLESVSRSPPYSHFNETLIGVSVIARAEDEBERIHOSSDLKVDENQAXYYSIVANRPLA 1200
QY 1201 VRLCEVGNCIYLPALFAVISRHSLSAGLVGLSVYSLOVTTYVNMIVRMSSEMETNIVA 1260
DB 1201 VRLCEVGNCIYLPALFAVISRHSLSAGLVGLSVYSLOVTTYVNMIVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPPSSWPQVRVEFRNYCLRYRDLDFLASHINVTINGG 1320
DB 1261 VERLKEYSETEKEAPWQIOETAPPSSWPQVRVEFRNYCLRYRDLDFLASHINVTINGG 1320
QY 1321 EKVGI VGTGKXSLTGLFRINESAGEIINGINIAKIGLHDLRKTITIIQDPPLF 1380
DB 1321 EKVGI VGTGKXSLTGLFRINESAGEIINGINIAKIGLHDLRKTITIIQDPPLF 1380
QY 1381 SGLRPMNDPFSQVSDSEWMTSLFAHLKDVSAALPDLDBECABGGENLSVGOROLVCL 1440
DB 1381 SGLRPMNDPFSQVSDSEWMTSLFAHLKDVSAALPDLDBECABGGENLSVGOROLVCL 1440
QY 1441 ARAALLRKTILVLDDEATAVVDLETTDDLQSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVL 1500

Db 1441 ARALLRKTILVLEBATAVPLETDDLIQSTIRTFQEDCTVLTIAHRLNTIMDYRVVL 1500
QY 1501 DKGEIQEYGASDILLQORGLFYMAKXAGLV 1531
Db 1501 DKGEIQEYGASDILLQORGLFYMAKXAGLV 1531

RESULT 12
ADB20865
ID ADB20865 standard; protein, 1531 AA.
XX
AC ADB20865;
XX
DT 20-NOV-2003 (first entry)
XX
DE MRP1 based cancer related protein SEQ ID NO:678.
XX
KM irinotecan; colorectal cancer; cervical cancer; gastric cancer;
XX lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
XX variant allele; multidrug resistance protein 1; MRP1; cytosolic.
OS Unidentified.
XX
PN W02003013533-A2.
XX
PD 20-FEB-2003.
XX
PF 23-JUL-2002; 2002WO-EP008200.
XX
PR 23-JUL-2001; 2001EP-00117608.
XX
PR 24-MAY-2002; 2002EP-00011710.
XX
PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI Heinrich G, Kerb R;
XX
DR WPI; 2003-354397/33.
XX
PT Use of irinotecan or its derivative for preparation of a pharmaceutical
PT composition for treating cancer in a subject having a genome with a
PT variant allele comprising a multidrug resistance protein 1
PT polynucleotide.
XX
PS Disclosure; SEQ ID NO 678; 100bp; English.
XX
XX The present invention describes a method for the use of irinotecan (I) or
CC its derivative for the preparation of a pharmaceutical composition for
CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic
CC cancer, or malignant glioma in a subject having a genome with a variant
CC allele which comprises a multidrug resistance protein 1 (MRP1)
CC polynucleotide (II). (I) has cytostatic activity. (II) or its derivative
CC can be used for the preparation of a pharmaceutical composition for
CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic
CC cancer, or malignant glioma in a subject, where the subject is a human
CC (preferably African or Asian) or a mouse. The present sequence represents
CC a sequence which is used in the exemplification of the present invention.
XX
SQ Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 6; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTWNTSNPDKCFQNTVYLWWPCFYLMACPFYFLYLSRH 60
Db 1 MALRGFCSADGSDPLMDMNTWNTSNPDKCFQNTVYLWWPCFYLMACPFYFLYLSRH 60
QY 61 DRGVIQMTPLNKTKTALGFLIMTYCWADLFYSPWERSRGITFLAVPLVSPILGITTLA 120
Db 61 DRGVIQMTPLNKTKTALGFLIMTYCWADLFYSPWERSRGITFLAVPLVSPILGITTLA 120
QY 121 TFLIQLERRKGVQSSGIMLTFWLVALVCAIALIISKINTALKEDAQVDLFRDITFYVYFS 180

Db 121 TFLIQLERRKGVQSSGIMLTFWLVALVCAIALIISKINTALKEDAQVDLFRDITFYVYFS 180
QY 181 LLLIQLVISCFSDBSPLFSETIHDNPNCPRESSASFLSRTITFWITGLIVRGYOPLEGSD 240
Db 181 LLLIQLVISCFSDBSPLFSETIHDNPNCPRESSASFLSRTITFWITGLIVRGYOPLEGSD 240
QY 241 LMSLNKEDTSROVYVPLVKNMKKECATRQKQPVVYVYSSKDPAPKSSSKDANBEVAL 300
Db 241 LMSLNKEDTSROVYVPLVKNMKKECATRQKQPVVYVYSSKDPAPKSSSKDANBEVAL 300
QY 301 IVKSPKEMNPSLFEVLYKTFGPYFLMSPFKALHDLMMFSGPOLIKLLIKFVNDTKAPD 360
Db 301 IVKSPKEMNPSLFEVLYKTFGPYFLMSPFKALHDLMMFSGPOLIKLLIKFVNDTKAPD 360
QY 361 WQGYFTVLLFVTACIQTVLHQYFHI CFSYGMRIKTAIVIGAVYRKALVITNSARKSSTV 420
Db 361 WQGYFTVLLFVTACIQTVLHQYFHI CFSYGMRIKTAIVIGAVYRKALVITNSARKSSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINMTWSAPLOYITALLYLMNLGFSYLVAGVAVMLMPVN 480
Db 421 GEIVNLSVDAQRFMDLATYINMTWSAPLOYITALLYLMNLGFSYLVAGVAVMLMPVN 480
QY 481 AYVAMKTKTYOVAAHMKSKDNRIKLMNELINGIKVLYAMELAPKDKVLAIROBELKYLK 540
Db 481 AYVAMKTKTYOVAAHMKSKDNRIKLMNELINGIKVLYAMELAPKDKVLAIROBELKYLK 540
QY 541 KSAVLSAVGTFTWVCTPFLVALCTFAVYVITDENNIIDAOFAFVSLFNLIRPEPLILP 600
Db 541 KSAVLSAVGTFTWVCTPFLVALCTFAVYVITDENNIIDAOFAFVSLFNLIRPEPLILP 600
QY 601 MVISIIQASVSLKRLKILFISHELEBDSIERRVYKGGGNSITVRNATFTWARSDDPT 660
Db 601 MVISIIQASVSLKRLKILFISHELEBDSIERRVYKGGGNSITVRNATFTWARSDDPT 660
QY 661 LINGITFSPGALVAVVGVGCGGKSLLSALLMEMDVBEHVAVIKGSVAVYPOQAMQND 720
Db 661 LINGITFSPGALVAVVGVGCGGKSLLSALLMEMDVBEHVAVIKGSVAVYPOQAMQND 720
QY 721 SLRENILFGCOLBEPYRSYIOACALPDLEILPSGDRTEIGEGVNLSSGQOKRVSILAR 780
Db 721 SLRENILFGCOLBEPYRSYIOACALPDLEILPSGDRTEIGEGVNLSSGQOKRVSILAR 780
QY 781 AVYSNADIYLPDPLSAVDNAHVGHKHEBNYIGPKMLKNTRIILYTHSMSTLPQVDYIIV 840
Db 781 AVYSNADIYLPDPLSAVDNAHVGHKHEBNYIGPKMLKNTRIILYTHSMSTLPQVDYIIV 840
QY 841 MSGGKISEMGSYOELARDGAPAFELRTYASTBOEDAEENGVTGVSQPGKEAKOMENGM 900
Db 841 MSGGKISEMGSYOELARDGAPAFELRTYASTBOEDAEENGVTGVSQPGKEAKOMENGM 900
QY 901 LVYDSAGKQOROLSSSSSYSGDISRRHNSYTAELQKAKEEFTKMEADKAQYGVKL 960
Db 901 LVYDSAGKQOROLSSSSSYSGDISRRHNSYTAELQKAKEEFTKMEADKAQYGVKL 960
QY 961 SVYWDYKAIQGLISFLSITLFCNHNVSALASNYWLSLWTDPIVNGTOEHTKYRSLVYG 1020
Db 961 SVYWDYKAIQGLISFLSITLFCNHNVSALASNYWLSLWTDPIVNGTOEHTKYRSLVYG 1020
QY 1021 ALGISOGIAVFGYSMAVSTGGIILASRCLAHVDLHSIIRSPSPFERNVLRFSKEL 1080
Db 1021 ALGISOGIAVFGYSMAVSTGGIILASRCLAHVDLHSIIRSPSPFERNVLRFSKEL 1080
QY 1081 DTVDSMIEVYIKMFMSLFFNVIGACIVILATPAAIIPPLGLIYFVORFYAASSROL 1140
Db 1081 DTVDSMIEVYIKMFMSLFFNVIGACIVILATPAAIIPPLGLIYFVORFYAASSROL 1140
QY 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEOBERFIHOSDLKVDENQAAVYPSIYANRWLA 1200
Db 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEOBERFIHOSDLKVDENQAAVYPSIYANRWLA 1200
QY 1201 VRLCVCNCIVLPAALFAVISRSLSGVLGVSVYSIQVTTYLNLVLRMSSEMETIIVA 1260

Db 1201 VRLECVGNCLVLFAPALPAVVISRHSLSAGLVGLSLVSYSIQVTTYLNLVVRMSSMETNIVA 1260
Qy 1261 VERLKEYSETEKEAPWQIOETAPPSSWPQVGRVREFRNYCLRYREDLDFVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKEAPWQIOETAPPSSWPQVGRVREFRNYCLRYREDLDFVLRHINVTINGG 1320
Qy 1321 EKVGIVGTGAGKSLTGLFRINSGAGEI11IDGINIAKIGLHDLRFKTI111PODPVLF 1380
Db 1321 EKVGIVGTGAGKSLTGLFRINSGAGEI11IDGINIAKIGLHDLRFKTI111PODPVLF 1380
Qy 1381 SGLSRMNLDPSSQVSDSEVMTSLSLAHLKDFVSALPDLDHFCAGGSENLVSQORQVCL 1440
Db 1381 SGLSRMNLDPSSQVSDSEVMTSLSLAHLKDFVSALPDLDHFCAGGSENLVSQORQVCL 1440
Qy 1441 ARALLRKTKI1VLDEATAVLDLETDLLIOSTIRTOFEDCTVLT1AHLNT1MDYTRV1VL 1500
Db 1441 ARALLRKTKI1VLDEATAVLDLETDLLIOSTIRTOFEDCTVLT1AHLNT1MDYTRV1VL 1500
Qy 1501 DKGEI0EYGAPSD1LQORGLFYSAKXAGLV 1531
Db 1501 DKGEI0EYGAPSD1LQORGLFYSAKXAGLV 1531

RESULT 13

ADB87954 standard; protein; 1531 AA.

ADB87954;

04-DEC-2003 (first entry)

Human UGT1A1 protein sequence SEQ ID NO:678.

irrinotecan; cancer; UGT1A1; cytostratic; topoisomerase I inhibitor;
KM colorectal cancer; cervical cancer; gastric cancer; lung cancer;
KM ovarian cancer; pancreatic cancer; malignant glioma;
KM uridine diphosphate glycosyltransferase1 member A1.

Homo sapiens.

WO2003013536-A2.

20-FEB-2003.

23-JUL-2002; 2002WO-BP008217.

23-JUL-2001; 2001EP-00117608.

24-MAY-2002; 2002EP-00011710.

(EPID-) EPIDAUROS BIOTECHNOLOGIE AG.

Heinrich G, Kerb R;

WPI; 2003-289896/28.

Use of irrinotecan to treat cancer patient by determining if patient has
variant alleles of UGT1A1 gene, administering increased/decreased amounts
of irrinotecan based on increased/decreased levels of UGT1A1 gene product.

Disclosure; SEQ ID NO 678; 107bp; English.

The invention relates to the novel use of irrinotecan to treat a patient
suffering from cancer. This involves determining if the patient has one
or more variant alleles of the UGT1A1 gene, and if the patient has one or
more of such variant alleles, irrinotecan is administered in an increased
or decreased amount in comparison to the amount that is administered
without regard to the patient's alleles in the UGT1A1 gene. The invention
has cytostratic activity. A composition of the invention acts as a
topoisomerase I inhibitor. The method is useful for treating a patient,
an animal e.g. mouse or a human, preferably African or Asian, suffering
from cancer such as colorectal, cervical, gastric cancer, lung, ovarian,
pancreatic cancer or malignant glioma. The present sequence is udes in
the exemplification of the invention.

XX Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 7; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRGFCSADGSDPLMDMNTVNTSNDFTKCFONTVLVWPCCYLLMACPEFFLYLSRH 60
Db 1 MALRGFCSADGSDPLMDMNTVNTSNDFTKCFONTVLVWPCCYLLMACPEFFLYLSRH 60
Qy 61 DRGYI0MTPLNKTKTALGFLLMTCVADLFYSFMRERGF1FLAVPLVSP1LLG1TTLA 120
Db 61 DRGYI0MTPLNKTKTALGFLLMTCVADLFYSFMRERGF1FLAVPLVSP1LLG1TTLA 120
Qy 121 TFLI0LERRKGV0SSGIMLTFMVLVCAALILSK3MTALKEDAOVDLFFDITFYVYS 180
Db 121 TFLI0LERRKGV0SSGIMLTFMVLVCAALILSK3MTALKEDAOVDLFFDITFYVYS 180
Qy 181 LLLI0LVLS0CFSDRSPLFSETIHDNPPCESSASFLSRITFWITGLIVRGYR0PLEGSD 240
Db 181 LLLI0LVLS0CFSDRSPLFSETIHDNPPCESSASFLSRITFWITGLIVRGYR0PLEGSD 240
Qy 241 LMSLNKEDTS0QVVPV1VKMKKECAKTRKQPVKVYSSKDPAQPKSSKYDANEVVAL 300
Db 241 LMSLNKEDTS0QVVPV1VKMKKECAKTRKQPVKVYSSKDPAQPKSSKYDANEVVAL 300
Qy 301 IVKSP0KEMNPSLKFVLYKTFGPFLMSFFPKA1HDLMMFSGPOLIKLLIFVNDTRAPD 360
Db 301 IVKSP0KEMNPSLKFVLYKTFGPFLMSFFPKA1HDLMMFSGPOLIKLLIFVNDTRAPD 360
Qy 361 WQGFYFVTLFVTA0CLOTVLH0YFHI0FVSGMR1KTAVIGAVYRKALVITNSARKSSTV 420
Db 361 WQGFYFVTLFVTA0CLOTVLH0YFHI0FVSGMR1KTAVIGAVYRKALVITNSARKSSTV 420
Qy 421 GEIYNLMSVDA0RPMDLATY1NMTWSAPLOYTALYLMLNGSVYLAGAVNMLMPVN 480
Db 421 GEIYNLMSVDA0RPMDLATY1NMTWSAPLOYTALYLMLNGSVYLAGAVNMLMPVN 480
Qy 481 AVMAKTKTYOVAMKSKDNRIK1MNEI1N1KYVLKYAMELAFDKVLA1ROBELKYLK 540
Db 481 AVMAKTKTYOVAMKSKDNRIK1MNEI1N1KYVLKYAMELAFDKVLA1ROBELKYLK 540
Qy 541 KSAVLSAVGFTWVCTPFLVALCTFAVYVITDENN1LDAQTAVSLAFN1LRPPLN1LP 600
Db 541 KSAVLSAVGFTWVCTPFLVALCTFAVYVITDENN1LDAQTAVSLAFN1LRPPLN1LP 600
Qy 601 MVISIYQASVSLKRLR1FLSHBELPDS1ERRPVKGGGNS1TVRNATFTWARSDDPT 660
Db 601 MVISIYQASVSLKRLR1FLSHBELPDS1ERRPVKGGGNS1TVRNATFTWARSDDPT 660
Qy 661 LMGITFES1PEGALVAVVG0VCGKSS1LSALLAEMDKVEGHVA1KGSAAVYPOQAM1OND 720
Db 661 LMGITFES1PEGALVAVVG0VCGKSS1LSALLAEMDKVEGHVA1KGSAAVYPOQAM1OND 720
Qy 721 SLRENI1LFG0CLEEPYRSV1OACAL1PDLLE1LPSGRTE1GEGKVNLSGG0K0RVSLAR 780
Db 721 SLRENI1LFG0CLEEPYRSV1OACAL1PDLLE1LPSGRTE1GEGKVNLSGG0K0RVSLAR 780
Qy 781 AVYSNAD1YLFDDPLSAVDHVGKH1FENV1GPKGMLKNKTR1L1VTHSM5YLPOVDV1YV 840
Db 781 AVYSNAD1YLFDDPLSAVDHVGKH1FENV1GPKGMLKNKTR1L1VTHSM5YLPOVDV1YV 840
Qy 841 MSGG1ISMGSY0ELBLADGAFAPFLRYASTBO0DAEENG1VGS0PGEAK0EMEGM 900
Db 841 MSGG1ISMGSY0ELBLADGAFAPFLRYASTBO0DAEENG1VGS0PGEAK0EMEGM 900
Qy 901 LVTHSAGK0LOR0SSSSYSYGD1SRHNS1YAE1OKAKKEETWKLMEADK0ACT0GYKL 960
Db 901 LVTHSAGK0LOR0SSSSYSYGD1SRHNS1YAE1OKAKKEETWKLMEADK0ACT0GYKL 960
Qy 961 SVYDYMKA1GL1FIS1LS1FLM0CNHVSALASNTW1SLMTDPIVNGT0EHTKVL5YVG 1020
Db 961 SVYDYMKA1GL1FIS1LS1FLM0CNHVSALASNTW1SLMTDPIVNGT0EHTKVL5YVG 1020

Db 961 SVYWDYKAIIGLFISFLSIFLMCNHVSALASNYWLSMTDPIVNGTOEHTKVRLSVYG 1020
QY 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLHSLIRSPMSFPERPESGNLVNRESEKL 1080
Db 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLHSLIRSPMSFPERPESGNLVNRESEKL 1080
QY 1081 DTVDMSIPEVIMKMGSLFNVIIGACIVILLATPIAIIIPPLGIYEFVQRFYVASSRQL 1140
Db 1081 DTVDMSIPEVIMKMGSLFNVIIGACIVILLATPIAIIIPPLGIYEFVQRFYVASSRQL 1140
QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANRWLA 1200
Db 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANRWLA 1200
QY 1201 VRLFCVNCICVLFALPAVIRSRHLSAGLVGLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260
Db 1201 VRLFCVNCICVLFALPAVIRSRHLSAGLVGLSVSYSLQVTTYLNWLVMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMOIQETAPPSWPQVGRVFRNVCLEFRDLDVFLSHINVTINGG 1320
Db 1261 VERLKEYSETEKEAPMOIQETAPPSWPQVGRVFRNVCLEFRDLDVFLSHINVTINGG 1320
QY 1321 EKVGIVGRTGAGKSLTLGLFRINESABGEIIDIINIAKIGLHDLRFKTIIPQDPVLF 1380
Db 1321 EKVGIVGRTGAGKSLTLGLFRINESABGEIIDIINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SSGIRNMLDPPSOYSDEEVTWLSLALHDKDFVSLPKDLHCEAGGENLSVGOVOLVCL 1440
Db 1381 SSGIRNMLDPPSOYSDEEVTWLSLALHDKDFVSLPKDLHCEAGGENLSVGOVOLVCL 1440
QY 1441 ARALLRRTKTLIVDEATAVADLETTDDLIQSTIRTOFEDCTVLTIAHRLNTIMDTRYIVL 1500
Db 1441 ARALLRRTKTLIVDEATAVADLETTDDLIQSTIRTOFEDCTVLTIAHRLNTIMDTRYIVL 1500
QY 1501 DKGEIQEYGAPSDLLQORGLFYSMAXDAGLV 1531
Db 1501 DKGEIQEYGAPSDLLQORGLFYSMAXDAGLV 1531

RESULT 14
ADB96937 standard; protein; 1531 AA.
XX ADB96937;
AC ADB96937;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human MDR1 related protein sequence SEQ ID NO:678.
XX
KM irinotecan; colorectal cancer; cervical cancer; gastric cancer;
XX lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
KM multidrug resistance 1; MDR1; cytosolic; human; Cyp3A5; MRP1; MDR1;
XX TDP1.
XX
OS Homo sapiens.
XX
PN W02003013537-A2.
XX
PD 20-FEB-2003.
XX
PF 23-JUL-2002; 2002W0-EP008218.
XX
PR 23-JUL-2001; 2001EP-00117608.
XX 24-MAY-2002; 2002EP-0001710.
XX
PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
PI Heinrich G, Kerb R;
XX
DR WPI; 2003-268145/26.
XX
PT New use of irinotecan for preparation of pharmaceutical compositions for
treating cancer in subject having genome with variant allele comprising

PT multidrug resistance 1 polynucleotide.
XX
PS Disclosure; SEQ ID NO 678; 130pp; English.
XX
CC The invention relates to the novel use of irinotecan or its derivative
CC for the preparation of pharmaceutical compositions for treating
CC colorectal, cervical, gastric, lung, ovarian or pancreatic cancer, or
CC malignant glioma in a subject having a genome with a variant allele which
CC comprises a multidrug resistance 1 (MDR1) polynucleotide. A composition
CC of the invention has cytotoxic activity. The invention is useful for the
CC preparation of pharmaceutical compositions for treating colorectal,
CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant
CC glioma in a subject (preferably human, more preferably African or Asian)
CC or a mouse. The present sequence is used in the exemplification of the
CC invention.
XX
SQ Sequence 1531 AA;
XX
Query Match 79.4%; Score 7860; DB 7; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSADGSDPLMDMNTWNTSNDFPKCFQNTVLYWVPCFYLMACEPFYLYLSRH 60
Db 1 MALRGFCSADGSDPLMDMNTWNTSNDFPKCFQNTVLYWVPCFYLMACEPFYLYLSRH 60
QY 61 DRGYQMTPLNKTALGFLMAYCADLYSTFERSRGIFLAPVFLVSPFLGITTLLA 120
Db 61 DRGYQMTPLNKTALGFLMAYCADLYSTFERSRGIFLAPVFLVSPFLGITTLLA 120
QY 121 TPLIOLERRKGVSSGIMLFWLVAVCALAIRSKIMTFLKEDAOVDLPDIFEFYVFS 180
Db 121 TPLIOLERRKGVSSGIMLFWLVAVCALAIRSKIMTFLKEDAOVDLPDIFEFYVFS 180
QY 181 LLLIOLVLSGFSRSPFLFSETIHDNPNCPRESSASFLSRITFWWTGILVGRYRPLEGSD 240
Db 181 LLLIOLVLSGFSRSPFLFSETIHDNPNCPRESSASFLSRITFWWTGILVGRYRPLEGSD 240
QY 241 LMSLNKEDISEOVVPLVKNWKKCAKTRQPKYVYSSODPAOPKSSRYDANEVEAL 300
Db 241 LMSLNKEDISEOVVPLVKNWKKCAKTRQPKYVYSSODPAOPKSSRYDANEVEAL 300
QY 301 IVKSPQKKNPSSLFKVLKTFPGPYPLMSFPFKAIHDLMPSSGQILKLLKFNVDTRAPD 360
Db 301 IVKSPQKKNPSSLFKVLKTFPGPYPLMSFPFKAIHDLMPSSGQILKLLKFNVDTRAPD 360
QY 361 WOGYFYTVLLFVTAQTLVLHOFPHICFVSGMRKIKTAVIGAVYRKALVITNSARKSSTV 420
Db 361 WOGYFYTVLLFVTAQTLVLHOFPHICFVSGMRKIKTAVIGAVYRKALVITNSARKSSTV 420
QY 421 GEIVNIMSVDAORPMDIATYINMISAPLOVIALYILMLNIGPSVLGAVAVWLVAVN 480
Db 421 GEIVNIMSVDAORPMDIATYINMISAPLOVIALYILMLNIGPSVLGAVAVWLVAVN 480
QY 481 ANVAMKTKQYOVAMSKONRITLMEILLNGITVLYKYAMELAFKOKVLAIRQEBELVKL 540
Db 481 ANVAMKTKQYOVAMSKONRITLMEILLNGITVLYKYAMELAFKOKVLAIRQEBELVKL 540
QY 541 KSAVLSAVGFTWCTPFLVALCTFAVYVYTIIDENNIIDAOTAFVSLALFNILRPLNILL 600
Db 541 KSAVLSAVGFTWCTPFLVALCTFAVYVYTIIDENNIIDAOTAFVSLALFNILRPLNILL 600
QY 601 MVISSIVQASVSLKRLRIFLSHELEPDSIERRPVDOGSGTNSITVYRNATFTWARSDDPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHELEPDSIERRPVDOGSGTNSITVYRNATFTWARSDDPT 660
QY 661 LNGITFSIPRGALVAVVGVGGKSSLSALALAMKVESHVAIKSVAVVPOQAMQND 720
Db 661 LNGITFSIPRGALVAVVGVGGKSSLSALALAMKVESHVAIKSVAVVPOQAMQND 720
QY 721 SLRENTILFGQLEPEPYRSYVIOACALLPDEILPDSGRTEIGKGVNLSCGQKORVSLAR 780
Db 721 SLRENTILFGQLEPEPYRSYVIOACALLPDEILPDSGRTEIGKGVNLSCGQKORVSLAR 780

QY 781 AVYSNADIIYLPDDPLSAVDAAVGHKII FENVIGPKGMLKNKTRILVTHSMSTYLPQDVYIIV 840
DB 781 AVYSNADIIYLPDDPLSAVDAAVGHKII FENVIGPKGMLKNKTRILVTHSMSTYLPQDVYIIV 840
QY 841 MSGGKISMSGSYOEILNADGAFAEFLRTYASTEQDABEENGVTYVSGPGKAKEMNGM 900
DB 841 MSGGKISMSGSYOEILNADGAFAEFLRTYASTEQDABEENGVTYVSGPGKAKEMNGM 900
QY 901 LVTSAGKQLOROISSSSSYSGDISRHHNSTAEILOKAKAKKETWKLMEADKAOYGQYKL 960
DB 901 LVTSAGKQLOROISSSSSYSGDISRHHNSTAEILOKAKAKKETWKLMEADKAOYGQYKL 960
QY 961 SVYWDYMAKIGLFLISFLSIFLMCMCHVSALASNYWLSMTDDPIYNGTOEHTKYALSYYG 1020
DB 961 SVYWDYMAKIGLFLISFLSIFLMCMCHVSALASNYWLSMTDDPIYNGTOEHTKYALSYYG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGGILASRCLHVDLHSHILSPMSFFERTPSGNLVNRFKEL 1080
DB 1021 ALGISOGIAVFGYSMAVSIIGGILASRCLHVDLHSHILSPMSFFERTPSGNLVNRFKEL 1080
QY 1081 DTVDSMTPEVIMKMFMSGLFENVIGACTIVLLATPIAIIIPPLGLIYFFVQRFYVASSHQL 1140
DB 1081 DTVDSMTPEVIMKMFMSGLFENVIGACTIVLLATPIAIIIPPLGLIYFFVQRFYVASSHQL 1140
QY 1141 KRLESVSRSPYVSHENETILGVSIVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLESVSRSPYVSHENETILGVSIVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLCEVGNCTVLFALPALPAVISRHSLSAGLVGVSYSIQVTTYLNMVLRMSSEMETNIVA 1260
DB 1201 VRLCEVGNCTVLFALPALPAVISRHSLSAGLVGVSYSIQVTTYLNMVLRMSSEMETNIVA 1260
QY 1261 VERLEKESYETEKBAFMQIOETAPSPSPQVGRVEFRNCLARYEDLDVLAHINVTINGG 1320
DB 1261 VERLEKESYETEKBAFMQIOETAPSPSPQVGRVEFRNCLARYEDLDVLAHINVTINGG 1320
QY 1321 EKVGIYGTGAGKSSLTGLFRINSAEGEIIIDGINIAKIGLHDLRFKTIIPDPVLYF 1380
DB 1321 EKVGIYGTGAGKSSLTGLFRINSAEGEIIIDGINIAKIGLHDLRFKTIIPDPVLYF 1380
QY 1381 SSGIRMANIDPPSQSDEDEWMTSLBLAKDPYSALPDKLHBCAGGENLSYGQOLVCL 1440
DB 1381 SSGIRMANIDPPSQSDEDEWMTSLBLAKDPYSALPDKLHBCAGGENLSYGQOLVCL 1440
QY 1441 ARALLRKTILVLEATAAVDLEETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYRVIVL 1500
DB 1441 ARALLRKTILVLEATAAVDLEETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYRVIVL 1500
QY 1501 DKGEIOEYGAPSDLLQGRGLFYSMADAGLV 1531
DB 1501 DKGEIOEYGAPSDLLQGRGLFYSMADAGLV 1531

RESULT 15
ADB92128
ID ADB92128 standard; protein; 1531 AA.
XX
AC ADB92128;
DT 04-DEC-2003 (first entry)
XX
Human MDRI related protein sequence SEQ ID NO: 678.
XX
irinotecan; colorectal cancer; cervical cancer; gastric cancer;
KM lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
KM multidrug resistance 1; MDRI; cytosolic; human; UGT1A1; MRP1; TOP1.
XX
OS Homo sapiens.
XX
PN WO2003013535-A2.
XX
PD 20-FEB-2003.

XX 23-JUL-2002; 2002WO-EP008220.
PF
XX 23-JUL-2001; 2001EP-00117608.
PR
XX 24-MAY-2002; 2002EP-00011710.
PR
XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
PA
XX Heinrich G, Kerb R;
PI
XX WPI; 2003-342400/32.
DR
XX
XX New use of irinotecan for preparation of pharmaceutical compositions for
PT treating cancer in subject having genome with variant allele comprising
PT multidrug resistance 1 polynucleotide.
XX
XX Disclosure; SEQ ID NO 678; 104pp; English.
PS
XX The invention relates to a novel use of irinotecan or its derivative for
CC the preparation of a pharmaceutical composition for treating colorectal,
CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant
CC glioma in a subject having a genome with a variant allele which comprises
CC a multidrug resistance 1 (MDRI) polynucleotide. A composition of the
CC invention has cytostatic activity. The present sequence is used in the
CC exemplification of the invention.

Sequence 1531 AA;

Query Match 79.4%; Score 7860; DB 7; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCADGSDPLMDMNVNTMTNSNPDFTKCFONTVLYWVPCFYLMACPFEFLYSRH 60
DB 1 MALRGFCADGSDPLMDMNVNTMTNSNPDFTKCFONTVLYWVPCFYLMACPFEFLYSRH 60
QY 61 DRGYIOMTPINKXTALGFLIMIVCMADLFYSFWERSRGIFLAEVPLVSPTLGITTLLA 120
DB 61 DRGYIOMTPINKXTALGFLIMIVCMADLFYSFWERSRGIFLAEVPLVSPTLGITTLLA 120
QY 121 TELLQERRKGVSSGIMLTFMWLVALCALAIIKSKIMTALKEDAOVDLFEDITFYVYS 180
DB 121 TELLQERRKGVSSGIMLTFMWLVALCALAIIKSKIMTALKEDAOVDLFEDITFYVYS 180
QY 181 LLLIQVLVSCPSDSSPIFSETIHDNPNCPRESSAFLSRITFMWITGLVIRGROPLEGSD 240
DB 181 LLLIQVLVSCPSDSSPIFSETIHDNPNCPRESSAFLSRITFMWITGLVIRGROPLEGSD 240
QY 241 LMSLNKEDTSBOVVPVLVKMKKCECAKTRKQPVKVVYSSKDPAPKSSKYDANEVEAL 300
DB 241 LMSLNKEDTSBOVVPVLVKMKKCECAKTRKQPVKVVYSSKDPAPKSSKYDANEVEAL 300
QY 301 IVKSPQKEMNPSLTKVLTKTGTGFPLMSFFPKAIIHDLMEGSPQILKLLIRFVNDTKAPD 360
DB 301 IVKSPQKEMNPSLTKVLTKTGTGFPLMSFFPKAIIHDLMEGSPQILKLLIRFVNDTKAPD 360
QY 361 WQGFYTYVLVFTACLOTLVTHQYFHIQFVSGMKIKRAVIAVARKALVITNSARKSTV 420
DB 361 WQGFYTYVLVFTACLOTLVTHQYFHIQFVSGMKIKRAVIAVARKALVITNSARKSTV 420
QY 421 GEIVNLSVDAQRPMDLATYINMWSAPLOVITLALYLLMLNGPSVLGAVAVMYLVN 480
DB 421 GEIVNLSVDAQRPMDLATYINMWSAPLOVITLALYLLMLNGPSVLGAVAVMYLVN 480
QY 481 AVNANKTKTYOVAAHKSNDKFIKLMNBIANGIKVLKUYAWELARDKYLAEROELKYLK 540
DB 481 AVNANKTKTYOVAAHKSNDKFIKLMNBIANGIKVLKUYAWELARDKYLAEROELKYLK 540
QY 541 KSAVLSAVGFTWCTPPLVALCTPAVYVTDENNIIIDAOAPFSLALFNILRPPLNTLP 600
DB 541 KSAVLSAVGFTWCTPPLVALCTPAVYVTDENNIIIDAOAPFSLALFNILRPPLNTLP 600
QY 601 MVISSIVQSVSLKRLRIFLSHBEI.BPDS.IERRPVKGGTNSITVRNATFTWARSDBPT 660

Db 601 MVISSIVQASVSLKRLIFLSHEELPDSIERRPVKGCGTNSITVRNATFTMARSDPPT 660
 QY 661 LNCITTSIPBEGALVAVVGVCCGKSSLSALLAEMDKGHHVAIKGSVAVYPOQAWIOND 720
 Db 661 LNCITTSIPBEGALVAVVGVCCGKSSLSALLAEMDKGHHVAIKGSVAVYPOQAWIOND 720
 QY 721 SLRENILFGCOLPEEPYRSVIOACALLPDLEILPSGDRTEIGEKVNLSSGQKORVSLAR 780
 Db 721 SLRENILFGCOLPEEPYRSVIOACALLPDLEILPSGDRTEIGEKVNLSSGQKORVSLAR 780
 QY 781 AVYSNADIIYLPDDPLSAVDHVGKHIFENVIGPKMLKNKTRILVTHSMSYLPQVDVILV 840
 Db 781 AVYSNADIIYLPDDPLSAVDHVGKHIFENVIGPKMLKNKTRILVTHSMSYLPQVDVILV 840
 QY 841 MSGGKISEMGSYQELILARDGAFAPFLRTYASTEOBDAENGVTGSGPKKAKOMENGM 900
 Db 841 MSGGKISEMGSYQELILARDGAFAPFLRTYASTEOBDAENGVTGSGPKKAKOMENGM 900
 QY 901 LVYDSAGKOLOROLSSSSSSSGDISRHHNSTAELOKAEKKEETWKLMEADKAOQGVKL 960
 Db 901 LVYDSAGKOLOROLSSSSSSSGDISRHHNSTAELOKAEKKEETWKLMEADKAOQGVKL 960
 QY 961 SVYWDYMKALGLFISPLSIFLFMCNHYSAALASNYMLSLMTDDPIVNGIOEHTKYRLSVYG 1020
 Db 961 SVYWDYMKALGLFISPLSIFLFMCNHYSAALASNYMLSLMTDDPIVNGIOEHTKYRLSVYG 1020
 QY 1021 ALGISQGIANFGYSMAVSIIGIILASRCLHVDLHSLRSFMSFFERTPSGNLVNRFSEKL 1080
 Db 1021 ALGISQGIANFGYSMAVSIIGIILASRCLHVDLHSLRSFMSFFERTPSGNLVNRFSEKL 1080
 QY 1081 DTVDMSMPEVIKXMGSLFNVIKACIVILATPTAATIIIPPLGLIYFFVQRFYASSROL 1140
 Db 1081 DTVDMSMPEVIKXMGSLFNVIKACIVILATPTAATIIIPPLGLIYFFVQRFYASSROL 1140
 QY 1141 KRLESVRSRPFYSHFNETLLGVSIVRAFEBOERFIHOSDLKVDENOKAYPSIVANRMLA 1200
 Db 1141 KRLESVRSRPFYSHFNETLLGVSIVRAFEBOERFIHOSDLKVDENOKAYPSIVANRMLA 1200
 QY 1201 VRLECVNGCIYLFALPAVISRHSLSAGLVGLSVYSLSQVTTYLNMLVRMSSEMETNIVA 1260
 Db 1201 VRLECVNGCIYLFALPAVISRHSLSAGLVGLSVYSLSQVTTYLNMLVRMSSEMETNIVA 1260
 QY 1261 VERLKESETEKEAPWQIOETAPSSWPQVGRVFRNYCLRYREDLDPVLRHINVTINGG 1320
 Db 1261 VERLKESETEKEAPWQIOETAPSSWPQVGRVFRNYCLRYREDLDPVLRHINVTINGG 1320
 QY 1321 EKYGIIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKITTIIIPQDPVLF 1380
 Db 1321 EKYGIIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKITTIIIPQDPVLF 1380
 QY 1381 SSGILRNMLDPPSOYSDEEWTSLLELAHLKDFVSALPDKLDHECAEGENLSVGQRQLVCL 1440
 Db 1381 SSGILRNMLDPPSOYSDEEWTSLLELAHLKDFVSALPDKLDHECAEGENLSVGQRQLVCL 1440
 QY 1441 ARALIRTKTILVDEATAVDLETFDDLIOGIRTOFEDCTVLTAAHRLNTIMDTYRIVVL 1500
 Db 1441 ARALIRTKTILVDEATAVDLETFDDLIOGIRTOFEDCTVLTAAHRLNTIMDTYRIVVL 1500
 QY 1501 DKGEIOEYGAAPSDILOQRLFYSMAXDAGIV 1531
 Db 1501 DKGEIOEYGAAPSDILOQRLFYSMAXDAGIV 1531

Search completed: December 15, 2005, 15:17:22
 Job time : 188.366 secs

This Page Blank (upside)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:57:04 ; Search time 37.8285 Seconds

(without alignments)
4901.320 Million cell updates/sec

Title: US-10-665-283-6

Perfect score: 9901

Sequence: 1 MALRGFCNSADGSDPLMDMNV.....ASVAVAKAKRPFSPDSLS 1927

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: PIR 80:*

1: pirl:*

2: pirl:*

3: pirl:*

4: pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	7849	79.3	1531	1	DVHUR	multidrug resistan
2	4484.5	45.3	1527	2	UE0336	canalicular multis
3	3587	36.2	1545	1	S71841	multidrug resistan
4	3507.5	35.4	1541	1	S71839	canalicular multis
5	3289	33.2	1494	2	E89447	protein P57C12.4 [
6	3225	32.6	1573	2	T21219	hypothetical prote
7	3195.5	32.3	1502	2	T42216	multidrug resistan
8	2838	28.7	1515	1	S51863	cadmium resistance
9	2732	27.6	1398	2	T20434	hypothetical prote
10	2548	25.7	1478	2	T38712	ABC transporter SP
11	2384.5	24.1	1623	2	T01369	ABC transporter At
12	2356.5	23.8	1495	2	D86428	glutathione S-conj
13	2326	23.5	1495	2	E86428	probable ABC trans
14	2294.5	23.4	1144	2	T27408	hypothetical prote
15	2263.5	22.9	1559	1	S64757	probable membrane
16	2238.5	22.6	1516	2	F86428	probable ABC trans
17	2230	22.5	1539	2	F84919	glutathione-conjug
18	2144.5	21.7	1355	2	T00961	ABC transporter-11
19	2144.5	21.7	1514	2	T52080	hypothetical prote
20	2140	21.6	1515	2	T52081	multi resistance p
21	2131.5	21.5	1490	2	T47840	MKP-like ABC trans
22	2129	21.5	1545	2	T46645	multi resistance p
23	2123	21.4	1153	2	T26883	hypothetical prote
24	2109	21.3	1545	2	T42751	hypothetical prote
25	2104.5	21.3	1511	2	T42711	sulfonamide recep
26	2098	20.9	1546	2	T42728	sulfonamide recep
27	2069.5	20.8	1661	2	T47796	ABC transporter-11
28	2062.5	20.8	1661	2	S64800	probable membrane

30	2037	20.6	1121	2	C87973	protein Y43P8C.12
31	2014.5	20.3	1582	2	A56248	sulfonamide recep
32	2008	20.3	1592	2	S48933	probable transport
33	1995	20.1	350	2	S68403	inward rectifier p
34	1980	20.0	390	2	JC4689	inwardly rectifyin
35	1923	19.4	330	2	A57616	inward rectifier K
36	1920	19.4	390	2	JC7901	inwardly-rectifyin
37	1888	19.1	1477	2	S64616	YOR1 protein - yea
38	1884.5	19.0	1037	2	T50518	ABC transporter-11
39	1805	18.2	1146	2	F84487	probable ABC trans
40	1722	17.4	946	1	JC5667	multidrug resistan
41	1693	17.1	1548	1	DVINS	multidrug resistan
42	1687	17.0	1421	2	T34225	hypothetical prote
43	1606	16.2	1427	2	T20903	hypothetical prote
44	1575.5	15.9	1427	2	T39219	acp-binding caset
45	1558	15.7	1469	2	T50210	probable ABC trans

ALIGNMENTS

RESULT 1

DVHUR
multidrug resistance protein (cell line H69AR) - human
N:Alternate names: multidrug resistance-associated protein (MRP)
C:Species: Homo sapiens (man)
C>Date: 30-Jun-1993 #sequence_revision 05-Dec-1998 #text_change 19-Jan-2001
C/Accession: A44231; A37495
R/Cole, S.P.C.; Bhargava, G.; Gerlach, J.H.; Mackie, J.E.; Grant, C.E.; Almquist, K.C.;
Science 258, 1650-1654, 1992
A>Title: Overexpression of a transporter gene in a multidrug-resistant human lung cancer
A/Reference number: A44231; MID:93088080; PMID:1360704
A/Accession: A44231
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: MAPRSGTGMGRGIPATPTSPARFRSSCGCLVFTSGPV, 50-1531 <COI>
A/Cross-references: UNIPARC:UPI00001746CB; GB:L05628; NID:91835658
A/Experimental source: Small cell lung carcinoma cell line H69AR
R/Cole, S.P.C.; Deeley, R.G.
Science 260, 879, 1993
A>Title: Multidrug resistance-associated protein: sequence correction.
A/Reference number: A37495; MID:93262415; PMID:8098549
A/Accession: A37495
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-60 <CO2>
A/Cross-references: UNIPARC:UPI00001746CC; GB:L05628; NID:91835658
A/Note: sequence extracted from NCBI backbone (NCBIP:131929)
C/Genetics:
A/Gene: GDB:MRP
A/Cross-references: GDB:136335; OMIM:158343
A/Map position: 16p13.1-16p13.1
C/Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C/Keywords: antibiotic resistance; ATP; duplication; nucleotide binding; P-loop; transmembrane; multidrug resistance; nucleotide-binding motif A (P-loop)
F/661-844/Domain: ATP-binding cassette homology <ABC1>
F/678-885/Region: nucleotide-binding motif A (P-loop)
F/788-792/Region: nucleotide-binding motif B
F/110-1503/Domain: ATP-binding cassette homology <ABC2>
F/1327-1334/Region: nucleotide-binding motif A (P-loop)
F/1450-1454/Region: nucleotide-binding motif B

Query Match 79.3%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	MALRGFCNSADGSDPLMDMNVNTSNPDPTKCFQNTLVWVPCFYLMACPFYLYSRH	60
DB	1	MALRGFCNSADGSDPLMDMNVNTSNPDPTKCFQNTLVWVPCFYLMACPFYLYSRH	60
QY	61	DRGYIOMTPLNKKTKTALGFLWIVCWADLFYFWSERSGIFLAPVFLVSGFTLLGITTLA	120
DB	61	DRGYIOMTPLNKKTKTALGFLWIVCWADLFYFWSERSGIFLAPVFLVSGFTLLGITTLA	120

QY 121 TFLIQLERRKGVSSGIMLTFMVALVYCALILRSKIMTALKEBAQVDLPFDITFYVYFS 180
Db 121 TFLIQLERRKGVSSGIMLTFMVALVYCALILRSKIMTALKEBAQVDLPFDITFYVYFS 180
QY 181 LLLIQLVLSGSDSPLEFSETIHPNCPRESSASFLSITFMWITGLIVRGYRQPLESD 240
Db 181 LLLIQLVLSGSDSPLEFSETIHPNCPRESSASFLSITFMWITGLIVRGYRQPLESD 240
QY 241 LMSINKEDTSQVAVPVLYKMKKCAKTRKQPVKVYVSSKDPAPKSSKVDABEVEAL 300
Db 241 LMSINKEDTSQVAVPVLYKMKKCAKTRKQPVKVYVSSKDPAPKSSKVDABEVEAL 300
QY 301 IVKSPKEMNBSLFKVLKYTFGPFYFLMSFFPKAIHDLMMFSGPOLIKLIRFVNDTKAPD 360
Db 301 IVKSPKEMNBSLFKVLKYTFGPFYFLMSFFPKAIHDLMMFSGPOLIKLIRFVNDTKAPD 360
QY 361 WQGFYTVLTFVTAQLOTLVHOFTHICFVSQMRKTAIVAGVYRKALVITNSARKSSTV 420
Db 361 WQGFYTVLTFVTAQLOTLVHOFTHICFVSQMRKTAIVAGVYRKALVITNSARKSSTV 420
QY 421 GEIYNLMSVDAQRFMDLATYINMISAPLOVILALYLMLNGPSVLGAVAVMLMPVYN 480
Db 421 GEIYNLMSVDAQRFMDLATYINMISAPLOVILALYLMLNGPSVLGAVAVMLMPVYN 480
QY 481 AVMAKTKTYOVAMHKSNDRIKLMEIILNGIKVLKLYAMELAFKDYALAIROBELKYLK 540
Db 481 AVMAKTKTYOVAMHKSNDRIKLMEIILNGIKVLKLYAMELAFKDYALAIROBELKYLK 540
QY 541 KSATLSAVGTFTWCTPFLVALCTPAVAVYVTDENNILDAQAFVSLALFNLRPPLNLP 600
Db 541 KSATLSAVGTFTWCTPFLVALCTPAVAVYVTDENNILDAQAFVSLALFNLRPPLNLP 600
QY 601 MVISIYQASVSLKRLRIFLSHEELPEPSIERRPVKDGGNSTIVRAVATFMARSDPT 660
Db 601 MVISIYQASVSLKRLRIFLSHEELPEPSIERRPVKDGGNSTIVRAVATFMARSDPT 660
QY 661 LMGITFSIPEGALVAVVQVCGKSSLLSALLAEMDKVEGHVAKGSVAVYVPOQAMIOND 720
Db 661 LMGITFSIPEGALVAVVQVCGKSSLLSALLAEMDKVEGHVAKGSVAVYVPOQAMIOND 720
QY 721 SLRENILFGCOLLEBYRHSVYQACALLPDLILPBGDTEIGEKEVNISSGQKORVSLAR 780
Db 721 SLRENILFGCOLLEBYRHSVYQACALLPDLILPBGDTEIGEKEVNISSGQKORVSLAR 780
QY 781 AVYENADLYLPDDPLSAVDAHVGKHIFFENVIQPKQMLKMKRILVTHSMTSLPOVDVITV 840
Db 781 AVYENADLYLPDDPLSAVDAHVGKHIFFENVIQPKQMLKMKRILVTHSMTSLPOVDVITV 840
QY 841 MSGKISMGSYOELLARDGAFABFLRTYASTBOQDAEENGVTGVSQPKGEAKOMENGM 900
Db 841 MSGKISMGSYOELLARDGAFABFLRTYASTBOQDAEENGVTGVSQPKGEAKOMENGM 900
QY 901 LVTHSAGQOLOROISSSSSYSGDISRHHNSTAELOKAKKBEYTKLMEADKAQOGYKL 960
Db 901 LVTHSAGQOLOROISSSSSYSGDISRHHNSTAELOKAKKBEYTKLMEADKAQOGYKL 960
QY 961 SVYDYMKAIGLIFSISIFLPMCHVASALSNWMLSLMTDDPIYNGTOEHTKURLSYG 1020
Db 961 SVYDYMKAIGLIFSISIFLPMCHVASALSNWMLSLMTDDPIYNGTOEHTKURLSYG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGLILASRCHVDLHSLILRSPMSFFERTPSGNLVNRFSEL 1080
Db 1021 ALGISOGIAVFGYSMAVSIIGLILASRCHVDLHSLILRSPMSFFERTPSGNLVNRFSEL 1080
QY 1081 DTVDSMTPEVIMKMGSLFNNIAGCTVILLATPIAIIIPGLGIYFVQRPYVASSQOL 1140
Db 1081 DTVDSMTPEVIMKMGSLFNNIAGCTVILLATPIAIIIPGLGIYFVQRPYVASSQOL 1140
QY 1141 KRLSVSVPYVSHNEFTLLGVSVIRAFEEQERFHOSDLKVDENOKAYVYSIVANRWLA 1200
Db 1141 KRLSVSVPYVSHNEFTLLGVSVIRAFEEQERFHOSDLKVDENOKAYVYSIVANRWLA 1200

QY 1201 VRLIECVNCIVLPAALFAVIRSRHSLAGVLGVSYSLOQTTYINMLVRMSSEMETNIVA 1260
Db 1201 VRLIECVNCIVLPAALFAVIRSRHSLAGVLGVSYSLOQTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMIOETAPSSWPQVGRVFRYRCIARYRBDLPVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKAPMIOETAPSSWPQVGRVFRYRCIARYRBDLPVLRHINVTINGG 1320
QY 1321 EKVGIIVGRTAGKSSLLTGLFRINESAGEHIIIDGINIAKIGLHDLRFKTIITIPQDPVL 1380
Db 1321 EKVGIIVGRTAGKSSLLTGLFRINESAGEHIIIDGINIAKIGLHDLRFKTIITIPQDPVL 1380
QY 1381 SGSIKMWLDPSSQYSDVEVWTSLELAHLKDPVSALPKLDHECAGEGENTSVGQROLVCL 1440
Db 1381 SGSIKMWLDPSSQYSDVEVWTSLELAHLKDPVSALPKLDHECAGEGENTSVGQROLVCL 1440
QY 1441 ARALLRRTKILVDEATAVADLETDDLIQSTIRIQFEDCTVLTARHLNTIMDTRYIVL 1500
Db 1441 ARALLRRTKILVDEATAVADLETDDLIQSTIRIQFEDCTVLTARHLNTIMDTRYIVL 1500
QY 1501 DKGEIOEYGA PSDLLOQRGLFYSMAKDAGLV 1531
Db 1501 DKGEIOEYGA PSDLLOQRGLFYSMAKDAGLV 1531

RESULT 2
JE0336
canalicular multispecific organic anion transporter - human
C:Species: Homo sapiens (man)
C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004
C:Accession: JE0336
R:Uchiiumi, T.; Hinoshita, E.; Haga, S.; Nakamura, T.; Tanaka, T.; Toh, S.; Furukawa, M.; Biochem. Biophys. Res. Commun. 252, 103-110, 1998
A:Title: Isolation of a novel human canalicular multispecific organic anion transporter, t.
A:Reference number: JE0336, MUID:99032812; PMID:9813153
A:Accession: JE0336
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1527 <UCH>
A:Cross-references: UNIPROT:O15438; UNIPARC:UP10000169859; GB:AF083552
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology C;Keywords: ATP
F:1306-1499/Domain: ATP-binding cassette homology <ABC2>

Query March 45.3%; Score 4484.5; DB 2; Length 1527;
Best Local Similarity 56.6%; Pred. No. 4.6e-296;
Matches 871; Conservative 261; Mismatches 372; Indels 35; Gaps 9;

QY 8 SADSDPLMDMNVMTWNTSNPDTKCFONTVYLWVPCFYLMACFPFYFLYLSRRHGRYIQM 67
Db 7 SGEIGSKFMDNLSVHTENPDLTFCFQNSILLAWPCILYLMVALPCYLLYLHHCGRYIIL 66
QY 68 TPLNKTITAGFLMIYICMADLFYSFWEBSRGITLAVVFLVSPILLGITTLATFLIQLE 127
Db 67 SHLSKLMVIGVLIMCVSMWDLFPYSFGLVGRAPAVFVFTPLPVVGVTMLATLTIQYE 126
QY 128 RRKVOSSGIMLTFMVALVYCALILRSKIMTALKEBAQVDLPFDITFYVYFSLLIQLV 187
Db 127 RLGVOSSGVLITWFLCVCAIVPERSKILDLAABGEISDPFFFTFYHFAVLVSLAL 186
QY 188 LSCFSDSPLEFSETIHPNCPRESSASFLSITFMWITGLIVRGYRQPLEGSDLSLNKE 247
Db 187 LACFREKPPFSACNVDPNPPERSAGFLSLFLFWMTKXALYGRHRLBEKDLMSLKEE 246
QY 248 DTSQVAVPVLYKMKKCAKTRKQPVKVYVSSKDPAPKSSKVDABEVEALIVKSPQK 307
Db 247 DRSQMVVQOLLEAARKQEKQYARH-----KASAPGK---NAGEDEVLLGAGPRP 294
QY 308 EWNLSLFLVLYKTEGPFYFLMSFFPKAIHDLMMFSGPOLIKLIRFVNDTKAPDQGYFT 367
Db 295 R-KPSFLKALLATGSSFLISACFKLIQDLISFLNPOLSTILIRFISNPMGSPWVGFLVA 353

QY	366	VLLPFTACLOTVLHQYEFICVSGMRITAVIAGVREALVITTSARKSSVGEHIVNM	427
Db	354	GLMFJCSMMQOJLHGYHYIFVTVGKFEFTGMYVRRFALVITNSVKKASVTGEBVINM	413
QY	428	SVDAORFMDLAFYINMTWSPARQVLIATLVLMTNLGSPVLAAVAVVVLVMPVNAVMAKT	487
Db	414	SYDARFMDLAFPLNLMNAPLOIILATIFLQNLGSPVLAAVAFVVLILPLANGAVAYVM	473
QY	488	KTYOVAHMKSKONRIKLMNEIINGIKVLKYAMELAFKQVLAIQOEBLKKVSAVISA	547
Db	474	RAFQYOMKLKOSRIKMEIINGIKVLKYAMEBSPFLKQVGBIRQOGLQRLTAAYLHT	533
QY	548	VCTFWVCTPPELVALCTPAVYVYTTIDENNILDKQTAFAVSLATNLIRPLNIPMTISIV	607
Db	534	TTTFPMWCSPELVYVITITLMMVYVVPNNVLDAEKAFVSVSLNNIIRLPLNMLPOLISMLT	593
QY	608	QASVSLKRLRIPLSHELEPDSIERBRPVYDGGTNSITVRNMTPTMAREDPPTLNGITFS	667
Db	594	QASVSLKRLQOFLSOEBLDPQSVERTKISPG--YAITHSSTFWAOOLPPLTHSLDIQ	650
QY	668	IPGALVAVVGVGGGKSLSLTALLAEMDKVEGHAIXGSVAVYVQOAMTQNDSDRENL	727
Db	651	VPKGALVAVVGVGGKSLVSGALLGEMEKLEKGVHMGGSVAVYVQOAMIQCTQENVL	710
QY	728	FGCQLEBPYRSVITQACALLPDLLEILPSGDRTEIGEGVNLSSGGKQORVSLARAVYSND	787
Db	711	FGKALNPKRYOQTLTACALLDLEMLPGDDQTEIGEKINISGGORORVSLARAVYSAD	770
QY	788	ITLPPDPLSAVANAHGKHFEENVICPKMKNKTRILVTHMSYLPQVDVITVMSGGKIS	847
Db	771	ITLDDPLSAVDSHAKHIFDHVIBPEGVLAQKTRVLVTHGISPLPQDPIITVLADGOVS	830
QY	848	EMGXYOELLARDGAFAEFLRTASTEOEDDAEENCVTSGSGPGEKAKOMENML-----	901
Db	831	EMGPRPALLQNRGSRANPLCNVAPBEDQCHLEDSTALEGADKALLIEDTLSNHTDIT	890
QY	902	---VTSAGKQJOLQOLSSSSSSSGD-----ISRHNSTAE-LQRAEAKKEBTWKLMEAD	951
Db	891	DNDPVTYVQKQFMKQSLMSS--DGEQQRVPVRHRLGSPSEKVQYTLAKAKDA--LTQBE	947
QY	952	KAQSGQVCLSVYMDMKALIGLFIPLSLFEMCNHVSALASVYVLSLWTDPIYNGTOEH	1011
Db	948	KAAITVLSVYMDAKAVGLCTTTLAICLLYVGOSAAAIKANVWLSAWTNAMDSDRON	1007
QY	1012	TKVRLSVGALGISOGIAVFGSVMSJGGLIASCJLVNDLHSLIRSPMSFEERTPSGN	1071
Db	1008	TSIRLGVYALGILQGLVLMANMAAGIOAAVVLQALLHANKIRSPQSFDTTSPGR	1067
QY	1072	LVNRESKELDTVDSMIPVYIKMFMSLFPNVIAGACIVILLATPEIAIIIPPLGLIFYFOR	1131
Db	1068	ILNCRSKOIIYVDEVILAPVILMLNSPFNAISTLVIVMASPTLPFVILLPLAVILTYLQ	1127
QY	1132	FYVASSROLKILIESVRSRPVISHNETLIGVSVIAPFEOBERFIHQSDIKYDENOKAYP	1191
Db	1128	FYAATSRQIKRLESVRSRPIYSHPESETVGSVIAVYRSDRFBELIISPTKYDANQSCYP	1187
QY	1192	SIVANRWTAALRECGNCIVLFAALFAVISHRSAGLVGJSVYSISLOVTTYLMLNVM	1251
Db	1188	YIISRMVLSIGEFEGNCVLPALFAVYIGKSLPGLVGVSVSISLOVTPALNMIMMM	1247
QY	1252	SEMENTIYAVERLKEVSETEKAPMOIOETAPPSMPOVGREFPNYCICLAREDDPVLAR	1311
Db	1248	SDLBENIYAVERVKEYSKTEHTEAPVYVGSRRPEBMPRGVREFPNYSVRRPRGLDYLK	1307
QY	1312	HINVTINGEKVGIYGRTGAGKSSITLGLFRINSEABEIIIDGINAKIGLHDLRFKIT	1377
Db	1308	DLISHVHGEKVGIVGRTGAGSSMTLCLFRLEAKAKEIRIDGNVADIGLHVRSQLT	1367
QY	1372	IIPQPVYFSGSLRNMLDPFSQYSPBEVWTSLEHLADPVSALDPDKDHECABEGNYS	1431
Db	1368	IIPQPIIFSGSLRNMLDPFGYSSEBDIKWALLESHLHTFPVSSQPAIGLDFQCSBEGNYS	1427
QY	1432	VGQROVLCLARALKKTLVLDEAYAAVDETDDLIOSTIRTOFEDCTVITIAHRLNTI	1491

Db 1428 VQORLVCLARPLLRKSKSLVLDSEATYAIIDLETNDLIDATYRTQPDCTVLTIAHRLNTY 1487

Qy 1492 MDYTRVILVDKGEIQEYCAPSDLLQOORGLFYSMARDAGI 1530

Db 1488 MDYTRVILVDKGVAAEFPSDPANLTIARCIFFGMARDAGI 1526

RESULT 3

71841

multidrug resistance protein, canalicular - human

C.Species: Homo sapiens (man)

C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001

C.Accession: S71841, S71840

R.Koenig, J.; Keppler, D.

A.Reference number: S71841

A.Accession: S71841

A.Molecule type: mRNA

A.Residues: 1-1545 <KOE>

A.Cross-references: UNIPARC:UPI00001746CD; EMBL:X96395; NID:g1507819; PIDN:CA65255.1;

R.Bucheler, M.; Koenig, J.; Brom, M.; Kartenebeck, J.; Spring, H.; Horie, T.; Keppler, D.

J. Biol. Chem. 271, 15091-15098, 1996

Article: cDNA cloning of the hepatocyte canalicular isoform of the multidrug resistance

A.Reference number: S71839; MUID:96279006; PMID:8662992

A.Accession: S71840

A.Status: nucleic acid sequence not shown

A.Molecule type: mRNA

A.Residues: 1415-1429, 'VP', 1432-1455, 'E', 1457-1545 <BUB>

A.Cross-references: UNIPARC:UPI00001746CE; EMBL:X96395

C.Genetics:

A.Gene: GDB:ABCC2; CMOAT; ABC; MRP2; cMRP; DJS

A.Cross-references: GDB:6089489; OMIM:601107

A.Map position: 10q24-10q24

C.Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

C.Keywords: ATP; glycoprotein; nucleotide binding; P-loop; transmembrane protein

F:651-837/Domain: ATP-binding cassette homology <ABC1>

F:671-678/Region: nucleotide-binding motif A (P-loop)

F:1317-1510/Domain: ATP-binding cassette homology <ABC2>

F:1334-1341/Region: nucleotide-binding motif A (P-loop)

Query March 36.2%; Score 3587; DB 1; Length 1545;

Best Local Similarity 47.5%; Pred. No.5.3e-235;

Matches 746; Conservative 286; Mismatches 461; Indels 78; Gaps 20;

Qy 3 LRGECSADGSDPLMDMVTWNTS-----NPDFIKCFONTVLVWPCFYLMACFPYFL-- 55

Db 2 LKFKCN-----STFWNSFLDSPDADPLCFCEGTVLVWIPLGFWLMLAPQGLHV 51

Qy 56 YLSHRDGYIOMTFLNKTKTALGFLMIVCAMDIFYSWEMSRGIFLAPVLVSTLLGI 115

Db 52 YKSTKTSSTTKLYLAK-QVVFGL-ILAIELALVLTESGQATPAVAVYVPSLY-L 108

Qy 116 TTLATFLIOLERRKGYOSSGIMLT-FMLVNLVCAALIRSKIMTALKEDAQVDLFRDT 174

Db 109 GTMLVLLIIQSRQWCQKSNWFLSLFIILISILCTGFQFLITLLIAGD-NSNIAYSL 167

Qy 175 FYVYFSLLLIQVLVSCFSDRSLPFSETHDNPCESSASFLSITFWMITGLVIRGYRQ 234

Db 168 PFISYQGIOLLIRFSARSEN-----NBSNNPSSIASFLSSITGYWYDSIILKGYR 219

Qy 235 PLBESDLMSLNKETSSQVPLVLRMKKECAKTRK-----QVYKVVYSKD 281

Db 220 PLTLEDWEVDEEKKTKTLVSKFETHMKRELQKRRALQRRQKSSQONSQARLPGLKN 279

Qy 282 PAQKESKVDANEVEVALIVKSPQKSWNP--LFKVLVYKFGCYFLMSFFKAIHIDLM 339

Db 280 QSGQSDALV---EDVEKKKKSGTKKDVPRSMKLKLFKFWVLKSLFLKLVDNIFT 336

Qy 340 FSGQIIKLILKFNVDKAPDMQGYFTYVLLFVTAACIQTIVLHQYFHCFSNGRIKTV 399

Db 337 FVSQILKLILISFASDSDYLVMLGYLCAIILFTALIQSFCLOCFYDLCFLGKVKRV 396


```
Db 344 GFVSSNSYVWFYICAILMFAVTLIIQSFCLQSYFOHCFVLGMCVRRTVMSIYKKALLTL 403
Qy 411 TNSRKSTVEIYNMVSVDQRFMDLATYINMWSAPLOVLIYLLMLAGSVLACV 470
Db 404 SNLRKRYTIGETIYNMVSVDQKLMADATNMVLSVIOITLIFLMLRGLPSIILAGV 463
Qy 471 AVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 530
Db 464 GAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAV 523
Qy 531 IROBELVYLKKSAYLSAVGFTVWCTPFLVALCTPAVVYVTDENNIIIDAQTAFFVALFN 590
Db 524 IRKKEKXNLRFGQLSILFILOITPILVSVFVSVVAVVAVVAVVAVVAVVAVVAVVAVV 583
Qy 591 ILRPLILPWNISIVQASVSLKRLRIFLSHELEPDSIRRPVKOGGNSITVRNAT 650
Db 584 ILRPLMLPWNISIVQASVSLKRLRIFLSHELEPDSIRRPVKOGGNSITVRNAT 639
Qy 651 PTMARSDPPTLNGITTSIPREGALVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 710
Db 640 PTMDDPELATIODVNLDIRKQALVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVVAVV 699
Qy 711 VPQAWIIONDSLRNIIIFGQLEBPYRSVIOACALLPDLIELSPGDRTEIGKGVNLG 770
Db 700 VPQSWIIONGTIKNIIIFGSEYNEKTYOVLKACALLPDLIELFGDMAEIGKGVNLG 759
Qy 771 GOKORVSLARAVVSNADIIYLEDPLSAVDHVGKHFENVIIGPKMLKNKTRIIIVTWS 830
Db 760 GOKORVSLARAAVYODADIYILDDPLSAVDHVGKHFENVIIGPKMLKNKTRIIIVT 819
Qy 831 YLPQVDVYIYMSGGKISEMGSYOBELARDGAFAEFLTYAS-TQEOBDAENGVYVSGP 889
Db 820 FLRPVDEIVVVGKTIIEKSGYRDLDDKGVFANMKTGKSGPSEBETVNN----- 872
Qy 890 GKEAKQWENQMLVT-----DSAG-----KOLQROLSSSSSSYGDSIRHNNSTAEIQ- 935
Db 873 DSEAEDEDDGLIPTMEIIPEDASLARMRENSLRITSSRSSRSRSGKSLAKSLKYNV 932
Qy 936 ---KAEAKKEBTWKLMADKAQTOGVLSVYWDYMKALGLEIFSLIFLFWCNHVASLAS 992
Db 933 NVLKEKEVEGQKLIKKEFEVETGKVFSTYKLYQAVGMMWSILFIILFGANNVARTGS 992
Qy 993 NYMLSLMT-DDPIVNGT---QEHKXRLSVYGAIGIOGIAVFGYSNAVSGIGLARCL 1048
Db 993 NLWMSAITSODNLTNGNSSSHDMKRIGVFAGLGAQIGICLLISTMSIYACENAKAL 1052
Qy 1049 HVDLHSLIRSPMSFFERTSGNLYNRFSEKELDVDSMIPEVIMFPGSLFNVIGACIVI 1108
Db 1053 HGQLLTNITLAPRPFPTTPTGRIVNRFSGDISIVDDLPLQTLRSMMWCFGIGITVMI 1112
Qy 1109 LLAITPAAIIIPPLGLIYFVQRFYVASSRQLKLEBSVSRSPVSHENETLLGVSVIRAF 1168
Db 1113 CMATPVPAAIIIPPLGLIYFVQRFYVASSRQLKLEBSVSRSPVSHENETLLGVSVIRAF 1172
Qy 1169 EBOERFTHQDLDKVDENQKAYPSIVANRMLAVRLEVCNICYLFAALPAVISHSISAG 1228
Db 1173 EHOERFTHQDLDKVDENQKAYPSIVANRMLAVRLEVCNICYLFAALPAVISHSISAG 1232
Qy 1229 LVGLSVSYLQVTTTYLWMLVRMSEMETNIVAVERLKEYSETEKEAAMQOIAETAPSWP 1288
Db 1233 VVGFVLSNALNITQTLWMLVRMSEMETNIVAVERLKEYSETEKEAAMQOIAETAPSWP 1291
Qy 1289 QVGRVEFRNYCLARYEDLDVLRHINVTINGEKVGIIVGTGAGKSLTLGLFRINESAE 1348
Db 1292 RHGEIIGNNYQVRRPELDVLRKIGITCNISGEKVGIVGRTGAGKSLTLGLFRINESAE 1351
Qy 1349 GEIIIGSINAKTIGHLRKRITIIPODPVLFGSSGLAMNIDPQOYDEDEWVMSLELAH 1408
Db 1352 GQIIBIDIVASIGHLRKRITIIPODPVLFGSSGLAMNIDPQOYDEDEWVMSLELAH 1411
Qy 1409 KDEVSALPDKLDBECAGEENTLVGQROLVCLARALARTKILVLDATAVAVDLETDLI 1468
```

```
Db 1412 RSPVSGQLGILLSVEEGDNLSIGRQLCLGRAVLRKSKILVLDATAVAVDLETDLI 1471
Qy 1469 QSTIRTFEDCTVLTIAHRLNTIMDYTRVVLVDGELIOEGABEDLLQORGLFYSMAKDA 1528
Db 1472 QTIIRKESQCTVITIAHRLHTIMDSKIMVLDNGKIVEGSPBELLNSRGSFYLAKEA 1531
Qy 1529 GL 1530
Db 1532 GI 1533

RESULT 5
E89447
protein F57C12.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: E89447
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.el
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an
A:Accession: E89447
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1494 <STO>
A:Cross-references: UNIPROT:Q20943; UNIPARC:UP1000017801E; GB:chr_X; PIDN:AAA83299.1; P
A:Gene: F57C12.4
A:Map position: X
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 33.2%; Score 3289; DB 2; Length 1494;
Best Local Similarity 47.6%; Pred. No. 9.3e-215;
Matches 696; Conservative 239; Mismatches 435; Indels 92; Gaps 18;

Qy 119 LAFLIQLERKGVQSSGIMLTFWLVAVCALMILSKINTALKEDAQVDLFRDITTYVY 178
Db 76 LAFLIYVACNKKIITSGVITLVLVVGQIEPFRYLSGFIYNEVALGIRATLYTIA 135
Qy 179 FSLILQLVLSGCSNDRPLSEITIHDPNCPRESSASPLRITFMWITGLVIRGYRPLEG 238
Db 136 FTSSALEFLCCPAD--VSDMYKSSSCPBETASTINLTITQWFTGLVYLGAKSLN 192
Qy 239 SDLMSINKEDTSQVVPVLYKNKKKCAKTRKQPVKVYSSKDP--AQPKSSKVDANEE 296
Db 193 EDLMDNEIDKAKNLIPISEFQNLKPRIDEVH-QNIK-----KQPSALPKN----- 237
Qy 297 VEALVYKSPQKWNPSLFFKLYTTFQGFYFLMSFFKRAIHDLMFSGFOILKLIKPVNDT 356
Db 238 -----HPSFVLPPIFTYKYTYLLAGFFYKLCFPMQLQFLAPQLKOLIGFTEDK 284
Qy 357 KAPDMQGYFTVTLVPTACQLTVLHQYFHICVSGMRITAVAGVYKALVITNSARK 416
Db 285 NQVWMTGCSIVGIMFSSFLQSFPLHQYHSMRILGHEVAVSITSAVYSRALNDSNARK 344
Qy 417 SSTVGEIVNMSVDAORFMDLATYINMWSAPLOVLIYLLMLAGSVLACVAVVAVM 476
Db 345 GKTIQAIYNMVSVDIQIDQIMAPMTLFWMSAPQIFLSTIFLWKFGLVAVLALAGVVLILA 404
Qy 477 VPVNAVMAMKTKTYQVAHMKSKONRIKLMNEIINGIKVTLKYVAMELAFKQVLAIROEL 536
Db 405 LPVNGIATIQMRKQTEOMKLDKERIKMSEIINGKMKVLCIYSWERSMEMVMVAKIRREL 464
Qy 537 KYVKSAYLSAVGFTVWCTPFLVALCTPAVVYVTDENNIIIDAQTAFFVALFN 595
Db 465 HILKLSYFMAAVFHWICAPFLASVYSFVYVYLLDENNVLTPELTVALSFLDLKMP 524
Qy 596 LNLIPWNISIVQASVSLKRLRIFLSHELEPDSIRRPVKOGGNSITVRNATFTWAR 655
Db 525 LAVAVAVVYGAVQGSVSNTRLKEFFAAEEMSPOT-----SISHGETDAIIEVNGLFPSWSS 580
```

```
QY 656 SDPEPTLNGITFSIPGALVAVVGVGCKSSLLSALLAEMDKVEGHVAKGSVAIVPQQA 715
DDEPTLREISFKIKGOLVAIVGKVGSKSSLLHALGEMNKLGSVQINGINIAVYVPOQA 640
QY 716 WIONDSIRENLFPGQLEEPYRSVIOACALLPLELIPSGDREIGKGNLSGGQKOR 775
641 WIONMSIRNNLLFPKPYDLENEYEDVKNKALKEDLANIPADREIGKGNLSGGQKOR 700
QY 776 VSLARAVYNSADIVLPDDPLSAVDAAVAKHIFENVIGPK-OMLNKKTITLTHSMSTYLPQ 834
701 VSLARAVYQNPDIILLDDPLSAVDSHVKGKIFENVISSSTGCKASKRIVLVTHGTLTKH 760
QY 835 VDVIIIVMSGKISEMGSYQELLADGAAFLRYASTQEPQDAEENGVTGVSQPG- 890
761 CDQILIVKEGTISELGTQOELLNNSGAFAEFLIEESKTRGRVASIGDGSVEDEIL 820
QY 891 KEAKOMENGMVTSAGKQLOROLSSSS-----SYGSDISR----HNSTAELOK 936
821 RDLQOVKPGIL-----KRLSHLSQESDKEDTSARAIEYGRDSSRSLHSPSQHEE 874
QY 937 AEA-----KKEETWKLMEADKAQGVYKLSVYMDYMKALGIFSLIFLPMCNH 987
875 NEALLGAISEDPVQENQGLEKETVEYKVKFVYIAVFOALISITPILTFEFLYVSSG 934
QY 988 SALASNTWLSLMTD-PIVNGTQETHKVRLSVYGALGISQGIAVFGYSMAVSGGILASR 1046
935 LGIISNFFYAKLSHAKSGNRRTSSDPAKMEIGYAVLWGQSGFVVLIASITIGVLRASR 994
QY 1047 CLHVDLHSLIRSMSPERFPGSGLVNRFSKELDTVSMIPBYIKMMSGLSPVIGACI 1106
995 ILHAGLGNIRKSPAPFDPVTPIGILNRIGIDEADRTLDPVIRHMSMTFENVATLV 1054
QY 1107 VILLATPAAIITIPPLGLIYFVQGFYVASSROLKRLSVSRSPVSHFNTELLGVSYIR 1166
1055 VIMWATPAGIAPALISIVYIVLAFYISTGRQLKLESASRSPYSHFQSIGASIR 1114
QY 1167 AFEBOERFIHOSDKVDENQKAYYPSIVANFMAVLEECVANCIVLFAALPAVISRHS- 1224
1115 AFGVADNFIKOSQORVDHLLIAYYPSIVANFMAVLEECVANCIVLFAALPAVISRHS- 1174
QY 1225 LSAGLVGSVSYSLQVTTYLMVIMVMSSEMETNIVAVERLKEYSTEEKAWQOETARP 1264
1175 LSAAGLVGSVSYSLQVTTYLMVIMVMSSEMETNIVAVERLKEYSTEEKAWQOETARP 1222
QY 1285 SSMQVGVGVEFRNYCLARREDLDFVLRHINVTINGEKGIVGRTGAKSSLLTGLFRIN 1344
1233 KSWBEKGEISIKNSVRRPGLDLVHGISAHIAESEKVGIVGRTGAKSSLLTGLFRIN 1292
QY 1345 ESARGEIITIDGINIAKIGLHDLREFKITIIPDPVLFSSGLMKNIDPFSQYSDEEYVTSLE 1404
1293 EADGGSIEIDGINIANIQLEQLRSCLTIVPDPVLFSGTMKNIDPFSQYSDEEYVTSLE 1352
QY 1405 LAHLKDPYSALPDLKDEHCARGENTL-----SVGQROLVCLARALLRKTKI 1450
1353 NAHLKDPYSALPDLKDEHCARGENTL-----SVGQROLVCLARALLRKTKI 1412
QY 1451 LVLDGAFAVADLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYRVVILDKGEIOEYGA 1510
1413 LVLDGAFAVADLETDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYRVVILDKGEIOEYGA 1472
QY 1511 PSDLIL-QQKGLFYSMADAGLV 1531
1473 PKNLLANPDGIFYSMAKDANVV 1494
Db
```

RESULT 6

T21219
hypothetical protein F21G4.2 - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T21219; T24002
R/Motifmore: B.
Submitted to the EMBL Data Library, October 1996

```
A/Reference number: Z19392
A/Accession: T21219
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1573 <M12>
A/Cross-References: UNIPROT:Q93552; UNIPARC:UPI00000823F7; EMBL:Z81016; PIDN:CAB02667.1
A/Experimental source: clone F21G4
R/Kershaw, J.
submitted to the EMBL Data Library, March 1995
A/Reference number: Z19828
A/Accession: T24002
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1573 <M12>
A/Cross-References: UNIPARC:UPI00000823F7; EMBL:Z48621; PIDN:CAA88549.1; GSPDB:GN00028; C/Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
Query Match 32.6%; Score 3225; DB 2; Length 1573;
Best Local Similarity 42.4%; Pred. No. 2.3e-210;
Matches 685; Conservative 305; Mismatches 490; Indels 134; Gaps 26;
QY 3 LRG---FCSADGSDPLMDMNTWNTSN-PDFTKCFQNTVAVWPCFYLMACFPYFYLS 58
9 LKGDAC-----GERYD-PAVWNASTVPLISQCYQHTLVMPF----- 47
Db
QY 59 RHDRGYIQMTPLNKTALGFLMIVCADLYSF-----WER-----SR 98
48 -----TAIVFLAPILTAQIFYRPNPDIWTRIRLOKIGLACILIASDL 91
Db
QY 99 GIFAIPF-----LVSEPTLIGITLALTEILOERRKGVSSGIMLFWIAL 146
92 SLFTVAIYETLFGQFPYAVDPVPLTCLAMVLTALIVSCRNIGYVSGSLFTSMVFT 151
Db
QY 147 VCALAIR-----SKIMTALKEDAQVDLPRDITFYVYFSLILLOLVSCFSRSP- 202
152 ISALPELLMTOQVINAEMAMWIDYPRCIAFTWFCCAFETYLHCYADSPGKYLS 211
Db
QY 203 HDNPPCESSASFSRITFWMITGLIVRGYRQPLESGDLSLANKEDTSEQVAVLVKWK 262
212 AARPSETTSSFLNRTIMWFNSLCSLGVKKPLEVSDIVSLNADTSNLLVPKVMILMD 271
QY 263 KECAKTKQPKVYVYSSKDRQPKSSKVDANEVEL-----IVKSPQKEMN 310
272 KQSKKFEETARRRIGSN--ASRTNRRRTSNDTTPLLNDOSTDDYGSVPAGQSTQK--M 327
QY 311 PSLFKVLYKTPGPFYFLMSFPFKALHDMFMSGPOILKLLIKFVNDTAKAPDMQGFYTVL 370
328 PSITWTFLMKMIVITAMFYKLSVDVLRPNPLLSLIRFTBELERPMQGVVLAFTM 387
Db
QY 371 FVTACLOTLVLAHQYFIIQVSGMRKIKTAVIGAYVRKALVITNSARKSTVGEIVNLMS 430
388 FFSALSLISLISHFYFLMYRGTGVQCLPAVYRKTLRLSNAARRKRETVGEIVNLMAID 447
Db
QY 431 AQRMDLATYINMTWSPAPLOVITLALYILMLNGSVSLAGVAVYMLPVAVYAMAKTKTY 490
448 VDRFOQITPQTMQYWSNPFOGLAFLLFOOLGVSVSGVAVWVLLPPIINFTIMIRKW 507
Db
QY 491 QVAMKSKDNRIKLMNELINGIKYKLYAMELAFKDKVILROBELKYLKKSAYLSAVGT 550
508 QIAQMTYKDEKTKMNVENVLNGIKYIKLYAMEPPMEQYIEDLREBELGILKKAALRTPSD 567
Db
QY 551 FTWCTPPLVALCTFAVYVITIDENNIIDAQFAVSLAFNLIRPPLNLPVVISIYQAS 610
568 MLNTASPFVVALSTFAPIFYIDPQNVLTPLAFAVSLTFNLQDRSPMGVALHITQVQV 627
QY 611 VSLKRLAIFSLHELEPDSIERRVKGGGNSITVKNATITW--AASDP-PTLNGITFS 667
628 VSNRLKEFLVSEELNVEAIDHR-ARD--NNDVICLEACLSMSASHQHPVPTLTINISFS 684
Db
```


[illegible]

C:Genetics: A:Gene: CESP:ED03G2.2
A:Map position: X
A:Introns: 73/3; 113/3; 176/1; 216/1; 258/2; 384/3; 470/3; 549/1; 644/3; 757/3; 764/2;
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 27.6%; Score 2732; DB 2; Length 1398;
Best Local Similarity 41.1%; Pred. No. 6.4e-177;
Matches 591; Conservative 287; Mismatches 472; Indels 88; Gaps 19;

QY 115 ITTLATFLIQLERRKGVSSGIMLTFWLVLVCALALRSKIMALKED--AQVDLPD 172
DB 27 LALLIATNEV---RRAGIHSSGPLECTIMLRAVAAPVEFYQMNTGSGPELVARIDFRY 83
QY 173 ITFVYFSLLIQLVLSCFSDRSLFSETIHPN----PCPSSASFLSRITFWITGL 227
DB 84 VAYLTFFPLVAEFVLHFVSDPPFM-----PRGYQLMKCPENANFISRLMLMFIQI 136
QY 228 IVRGROQLBESDLSLNKEDTSEQVYVVLVKNKKCAKTRKQPVK-----VYSS 279
DB 137 ISLGERFLVADVDPEMDSQMDQELKAKMTKTEMLKQTEKAREKQVKLDDKRRARTGE 196
QY 280 KDPAPQKSSKYDAN--EEVEALIVKSPQKWNPSLFLKYLITGFGFYPLMSFFPKAIHIDM 338
DB 197 KAPLLGTFNNYGAVALDDKRVIVQ-----PSVITLMQIMKWEILGGSFIKLSDL 249
QY 339 MFSGPQIILKLIRKVNDRKAPDMQGYFTVLLFYVALCLOTILVHQFHCIFVSGMRKITA 398
DB 250 QFAPPTFLNVLILFLFPNAPLINGIGLAVGLFLAGQIKSLFPMNTYFIAMTRVGAKIQTM 309
QY 399 VIGAVYKALVITNSARKSSTVGEIVNLMSVDAQRFMDLATYINNTSAPLQVITIALYLL 458
DB 310 LSCAVYESKSLLSNTARERTVGEMVNILSIDVRFMTPIQIQYWSSPQIILICVILL 369
QY 459 WLNIGPSTLVGAVMVLVYNVNAVMAMKTQYOVAHMKSKNRRIKLMEIINGITVLTLY 518
DB 370 SQTIGVANWAGIIVMISIVPINICVSMITTKWQILMKYKQERIRLNEVINGIKVILS 429
QY 519 AMELAFQDKVLAIQOEELKVLKKSAYLSAVSTFTWVCPTPLVALCTFAVYTYIDENIILD 578
DB 430 AMERNMEETIERVADKELKMKIKQSALIKTFADCLNAGAPVVALSFFVFLIDPKAVLT 489
QY 579 AQTAFFVSLATFNILRFPNLIPMVIVISSIVQASVSLKRLIFLSHEELEPDSIERRPVKG 638
DB 490 PNIAFVSLTFNLIRGLPMAAEFLVAQTVQLVNSKRRTFLCEKEVDTAIDKE--IRGE 548
QY 639 GGNISITVRNATFTMAASDPTLNGITFSIEGALVAVVGVGVGKSSLLSALLAEMKY 698
DB 549 LYTTFVIEHSGSFAMDSEARILSDIDFLAGSKELVTVVGVGVGKSSLLAALGEMKV 608
QY 699 EGHVAIGSVAYVYVQQAAMIONDSLRENILFPCOLEEPPYRVRVIOACALLPLEILPSGDR 758
DB 609 CGYGVGVSAYVLSQOPIILNOSLKNVLMQADINDVLYKIVISSCALKEBKLQLPDGD 668
QY 759 TELGEKGVNLSGGQKQKQVSLARAVYSNADITLFPDPLSAVDAAHYGKHIFENVIGPKMLK 818
DB 669 TEIEKGGINISGGQKARIALARAVYQSKDYFLDDPLSAVDAAHYGKHIFENVIGPKMLK 728
QY 819 NKRITLVTHSKSYLPQVDVITVMSGGKISMGKSYQELLARGAFAELRTYASTBOGDA 878
DB 729 HTTILITNCSFPLQESGKIIVMKG-----ERRYESG 761
QY 879 EENVVTGVSQGKRAKQEMNGMLVTDSAGKQORQLSSSSYSQDISRHHNSTAELQAE 938
DB 762 EESG-----GEENSDFLPQSIASGSRMSFLR--LSK-----ISKKSXSSTIVER-- 804
QY 939 AKKEETKTLMEADRAQQTGVQLSVYMDYKRAIGLIFSLSLFLPMCHVS--ALASNYVS 997
DB 805 -KKQDA--LITKEBAALGRVNPQVYLLYFKAMGIVTVYVLPALVAVLVNVSFALGRSLWLT 861
QY 998 LMTDDPLVNGTQEHKKY--RLSYGALGISQGLAVFGSAVNSVGGIILASGCHVDLHS 1055
DB 862 AMSAANDIDNHPTMSVGARLVAGGIGIEVILFFSLVYLLLGVAASNLKHKPLHAN 921

[illegible][illegible]

Db 1232 PSSGSIFEDVLAHYREDELPEVLHGVSFLSPMDKVGIVGRTGAGKSLNALFRITYLE 1291
Qy 1348 EGEIINDGINIAKIGLHDAFKITITIPDPLFSGSIRMLDPPSOYSDERWTSLELAH 1407
Db 1292 KGRLLIDECIGIRGLMDLKRVLGIIIPQAPVLFSGYRPNLDPFSEHNDLWESLRAH 1351
Qy 1408 LKDFVSLPDKLDECAEGENLSVGOROLVCARALLRKTLVLDEATPAVDLETDDL 1467
Db 1352 LKOTIRNPLGLDPEVTEAGENSVGOROLSLARLRSKTLVLDEATPAVDRTDL 1411
Qy 1468 IOSTIRTOFEDCVTLTAHRLNTIMDYTRVILVDKGEIOEYGAPSDLLQOGLFYSNAKD 1527
Db 1412 IQOTIREFESCTMLIIAHLNLTIIDCDKVLVDLSKVGOFESSPENLISNGESFSK--- 1468
Qy 1528 AGLVGGGGGMLSRKGIIPBEYVITRLAEDPAEPRYTRE 1567
Db 1469 -----MVOSTGTANAELRSITLENK-----RTRE 1493
RESULT 13
E86428
Probable ABC transporter [Imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: E86428
R/Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
Aneen, N.F.; Hughes, B.; Hlizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maizel, R.; Matzall,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: E86428
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-1495 <STO>
A/Cross-references: UNIPROT:Q9C8H0; UNIPARC:UPI0000482FF; GB:AE005172; NID:G1055818; F
C/Genetics:
A/Map position: 1
C/Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
Query Match 23.5%; Score 2326; DB 2; Length 1495;
Best Local Similarity 34.3%; Pred. No. 3e-149;
Matches 541; Conservative 314; Mismatches 566; Indels 154; Gaps 28;
Qy 2 ALMGFCS--ADGSDPLMDNVTNNTSNDFTCFQNTVLWVP-----CFY-LMACFP 51
Db 5 ALMWYCPVADGF-----WEKAVDGAFAVTPCAIDSLVMLVSHFVLLGLCFRIWITF- 58
Qy 52 FYFLVLSRHRRGVIQMTPLNKTALGFLIMVCMAD-----LFSFMRSRGIFLAPV 105
Db 59 -----HNTKAOIYVLRKKYCNVIG-LIACYCVAEPVRLVWGSLFPMDEETDPFP 110
Qy 106 FLVSPTLIGITTLATFLQLERKRGVSSGIMLTFVLVLCALMIRSKINTALKEA 165
Db 111 EVASLWVEAFAMFMSMLVIGLETKQYKEFRWYVRGVLYVADAVLDLVLPKRN 170
Qy 166 QVDFRDT--FYVYFSLILQLVLSGFSRSPPLFSETIHP-----NPPC 209
Db 171 RTLVYFISGRCSQALFGILLIYIPEL--DYPGYHIVNNEPLDNEVDALRGERICP 228
Qy 210 ESSASFSLRTFMWITGLYRGROPLESGDLSLWLNKEDPSQOVPLVNMKKECAKTR 269
Db 229 ERHASFSLRTYFMWITGLYRGROPLESGDLSLWLNKEDPSQOVPLVNMKKECAKTR 286
Qy 270 KQPVKVVYSKDDPAOPKSSKVDANEVEALIVYSPKEMNPSLKFVLYTFPGPYFLMSF 329
Db 287 -----PR-----PVLRLALNNSLGRFMTLAG 307

Qy 330 FKAIDLMWFGSPQILKLIKFNVDTKAPDMQGFYTVLLFYTACQLTVLHQYHICF 389
Db 308 IFKIGNDLSQFVGPVLLSHLRSMQED--PAMGVYAFILFPGVTLVGCBAQYFQNVW 366
Qy 390 VSGMRKTAIVIGAVYRKALVITNSARKSSTVGEIVMLMSYDAORFMDLATYINMWSAPL 449
Db 367 RVGFRLRSTLVAALFHKSLRLTHEARKNPFASGVYNTMITTDANAALQOISQHLGWAPF 426
Qy 450 QVILALYLLMLNIGPSVLGAVAVMIVPVAVAMAKTKTYQVAHMSKDNRIKLMNEIL 509
Db 427 RIIVSMILLYQQGVASLFSGLILFLIPLQTLIIKMRKLTKEGLQWTDKRVGIVTBEIL 486
Qy 510 NGIKVLKLYWELAFDKVLAIQEELKYLKSAVYLSAVTFTWCTPFLVACTPAVYV 569
Db 487 ISSMDTVKCYAMWESFSRIGINNEELSWERKQLLSAFNSFLINSLPVAVTVVSSFCV 546
Qy 570 TIDENNLDAQTAFFSIALFNILRPLNLIPLWYISSIVQASVSLKRL-RIFLSHELEDP 628
Db 547 LIGGD--LTPARAFSLSLFAVIRFPLNMLPULLSOVYANAVSLQRIEELLSEBRILAQ 604
Qy 629 SIERRPVKGGGNTSITVRNATFTW--ARSDPPLNGITFSIPGALVAVVGVQCCGKSL 687
Db 605 NPLQF-----GTPALSIKNGVSMDSKTKPTLSIDINIEIPVGLVAVIGTGEGKTSL 659
Qy 688 LSLALLAEMDKVE-GHVAIKSVAVYVQOAMTQDSLENTLRECGQLEBPYRYIOACAL 746
Db 660 ISMLDELSHAETTSVIRSVIRSVYVPOVSVIFATVARENTLFGSDDESERYWDAIDTAL 719
Qy 747 LPLELILPSGDRTEIGEKGNLSGGOKORSILARAYSNADILPDPPLSADVAHAKHI 806
Db 720 QHDLDLPLGRLDTEIGRGVNSIGGOKORSMAKAYNSDVITFDPLSALDAHVAHQV 779
Qy 807 FENVIPKGMKNKRTIILVTHSMYLPDVVDIIVMSGKISENGSOELIARDGAFAPFL 866
Db 780 FDSGM--KDELKGTHTVLVNTQHLPLMDXIKILVEBGMKEGTVELSKSLIFKGM 837
Qy 867 RITYASTEBQDAEENGVTGSGPGKEAKOMENMLYTDAGKOLQOLSSSSSGDISR 926
Db 838 ENAGKMDATQEVNTN-----DENLIKGLPTTVVDS--ERNLGSTK----- 876
Qy 927 HNSTALEKAEKKEETWMEADRAQGOVLSYWDPMKAI-GLFISFLIFPMCN 985
Db 877 -----QKRRRS-VLTKOBERETGIIISWNLARKYKAVGVGMVIMILACYLAT 924
Qy 986 HVSALASNYWLSLMTDPIYNGTQEHK-----VRSYVYGALGISOGIAVFGYMAVSIG 1040
Db 925 EVLRVSSSTWLSITWD-----QSTSKYVSPGYIVVYVVALGFGQVAVFTNMFMLITS 977
Qy 1041 GILASCLHVDLHLSLRGSMSPFERTPGNVLNRSKELDTVDSMIPEYIKMFGSLFN 1100
Db 978 SLHAARRLHDMSSILRAPMLFFHTNPTGRVNRFSKDIGIDRVANAMIMNFQMLQ 1037
Qy 1101 VIGACVILLATPIAIIIPRLGIYFPVQRFVYASSROLKULESRSRVSYHFMETLL 1160
Db 1038 LSTFALIGVSTISLWAIMPLILFYAALVLYQSTSRERRADSVTRSPIYAQFEBALN 1097
Qy 1161 GVSIVAPAEBOERFIHQSDLKVDENOKAYYSIVARWMLAVRLEGNCIVLPAALFAVI 1220
Db 1098 GLSSTIAYKAYADMAKINKSKMDNIRFLTANSSNRMLTIRIETIGGWIMTLTAFFAVL 1157
Qy 1221 SRHSLS-----AGVLGVSVYSYLOVTVYVNLVMSSEMETNIVAVERLKEYSETEKAP 1275
Db 1158 QNGNTNNOAGFASGTWGLISYTLINTISLSGLVLRQASRAENSINSVERVANYIDLSEAT 1217
Qy 1276 MOIQETAPSSPQVGRVVEFRNYCLRYREDLPVLVHINVTINGCKGVIGRTGAGKS 1335
Db 1218 DIIENNRPVCGMPPGSGSIFEDVHLRPGCLPVLGLTFYVSPSEKVGAVGRTGAGKS 1277
Qy 1336 LTLGLFRINESAGEIINDGINIAKIGLHDAFKITITIPDPLFSGSIRMLDPPSOY 1395
Db 1278 MNALFRIVEVEGRIMDDCVAKRGLTDVRAVSLTIQSPVLFSGYRPNLDPFSEHN 1337

A:Residues: 1-255 <BOY>
A:Cross-references: UNIPARC:UPI00001746CF
A:Note: the authors translated the codon CAG for residue 248 as His
R:Purnelle, B.; Goffeau, A.
Submitted to the EMBL Data Library, April 1996
A:Description: The sequence of 32 kb on the left arm of yeast chromosome XII reveals 14
mily and a new ABC transporter homologous to the human multidrug resistance protein.
A:Reference number: S69380
A:Accession: S69391
A:Molecule type: DNA
A:Residues: 1-1559 <PDR>
A:Cross-references: UNIPARC:UPI0000126A9C; EMBL:X97560; NID:g1297003; PIDN:CAA62776.1; F
R:Mioga, T.; Zimmermann, F.K.
Yeast 12, 693-708, 1996
A:Title: Sequence analysis of the CEN12 region of Saccharomyces cerevisiae on a 43.7 kb
e conductance regulator protein CTR.
A:Reference number: S70557; MUID:96405918; PMID:8810043
A:Accession: S70560
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1559 <MW>
A:Cross-references: UNIPARC:UPI0000126A9C; EMBL:X91488; NID:g1495203; PIDN:CAA62776.1; F
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1995
C:Genetics:
A:Gene: SGD:BPT1; MIPS:YLL015W
A:Cross-references: SGD:S0003938
A:Map position: 12L
A:Superfamily: human multidrug resistance protein CMOAT2; ATP-binding cassette homology
C:Keywords: ATP; nucleotide binding; P-loop; transmembrane protein
F:28-44/Domain: transmembrane #status predicted <TM1>
F:143-159/Domain: transmembrane #status predicted <TM2>
F:178-194/Domain: transmembrane #status predicted <TM3>
F:334-350/Domain: transmembrane #status predicted <TM4>
F:421-437/Domain: transmembrane #status predicted <TM5>
F:526-542/Domain: transmembrane #status predicted <TM6>
F:550-566/Domain: transmembrane #status predicted <TM7>
F:654-667/Domain: ATP-binding cassette homology <ABC1>
F:672-679/Region: nucleotide-binding motif A (P-loop)
F:974-990/Domain: transmembrane #status predicted <TM8>
F:1017-1033/Domain: transmembrane #status predicted <TM9>
F:1099-1115/Domain: transmembrane #status predicted <TM10>
F:1118-1134/Domain: transmembrane #status predicted <TM11>
F:1132-1128/Domain: transmembrane #status predicted <TM12>
F:1119-1522/Domain: ATP-binding cassette homology <ABC2>
F:1336-1343/Region: nucleotide-binding motif A (P-loop)
Query Match 23.2%; Score 2294.5; DB 1; Length 1559;
Best Local Similarity 35.1%; Pred. No. 4.4e-147;
Matches 574; Conservative 292; Mismatches 541; Indels 227; Gaps 39;
QY 23 NTSNPDFTKCFONTLVVWPCFYLM-ACFPYFLYLSRHDGVIQMTPLNKTATLGL- 80
DB 23 NALNP-----CFISVISAMQAVFLLIGSYQLMKLYKNNKVPFRKKNPPTLPSKINSRHLT 78
QY 81 -LWIVCW-----ADLFYSFMRSGIFLAPVLSPTLLGI---TITLAT 121
DB 79 HLTNVCQSTLIICELALVSSSDRYVPFLIK-KALYLNLF-----NLGISLPTQYLA 132
QY 122 FLIQLERRKVOSSGIMLTFWLVALVCAALIRSKI-----MTALKDAOV 167
DB 133 F-----KSTFSNGNDLFYFMPOILLQFLILQRYHSGSNERYLTVISQGTAMILEVL 185
QY 168 DLPRDITFYVFSLLIQLVLSCFSDRSPLFSETIHDNPNCPRESSASFLSRITFWITGL 227
DB 186 -LFPNSVAFIY-DLCIFEPI-----NELSEYKKNKWP-----PVHLSYITFIWANKL 233
QY 228 IYRGYRQPLRGSDLSLNLK--EDTSEQVAPVLVGNMKKECAKTRKQVKKVYSSKDAOP 285
DB 234 IVEYTR-----NKKIKDPNQPLPPVDLNK----- 259
QY 286 KESSKVDANEVEALIVKSPQKWNPSLFKVLYTFTPPYFLMSFFFAIDLMFMFSGPOI 345
DB 260 SISKEFKANWELKMLNRN-----SLMRAIWKSPRTISVAMLYETTSDLISLVQPOF 312

QY 346 LKLLIKFVN---DTKAPDMQGYFTVLLFTYACQLTVLHQYHICFVSGMRITAVIGA 402
DB 313 LRIPIIDGALPBTSSKYPPLNGVFTALTFLVISVSVFTNQYIGITGAGIGINGSLAST 372
QY 403 VYRKALVITVSARSKSTVGEIVNLMSYDA---GRFMDLATYINMWAPLOVIALYLLW 459
DB 373 VYQKSLATTLAERNEKSTGDLINMSDVLRIGQFFENA---QTLIGAPIDIIYVLSLY 429
QY 460 LNIQPSYLAGVAVNVLVPPVNAVAMKTKTYQVAHMSKDNRIKLMNEIINGIKVLYA 519
DB 430 WLIGKAVIGGLVWMAIMNPINAFLSRKVKLSKTQMKYKDMRIKTIETELNARIKLYA 489
QY 520 WELAFKRVLAIRE-ELTKVLSKASVLSAVGTFTWCTPPLVALCTAFAVYTTIDENNLD 578
DB 490 WEEPMARLHVNRDMLKNFRKGIYSNLIYFAMNCVPLMWCTSTGLF-SLPSDPLS 548
QY 579 AQTAFAVSLATFNILRFPNLILPMVISIVQASVSLKRLIFLSHEELEPDSIER-RPVKD 637
DB 549 PAIVFPLSLPNIILNSAIYVPSMINTIETSVSMERLKSFLSDEIDDSFTERIDPSAD 608
QY 638 GCGTNSITVNAATFTW-----ARSDPTLANGIT-----FSIPEGALV 674
DB 609 ERALPALEMNNITFLMKSKEVLTSSQGDMLRTDEESIIIGSSQALAKNIDHFEAKRGDLY 668
QY 675 AAVGQVCCGKSSLLSALLAEMDKVEGH-----VAIKSVAYVPOQAMIQNDSLKNT 726
DB 669 CVGRVAGKSTFLKALIGQLPCMSGSRDIPKPLIIRSSVAACSOESWIMNVSRENI 728
QY 727 LFGQLEPYPYRSYVIOCALLPLEILPSPDRTEIGKGVNLSCGQKORVSLAFAVYNSA 786
DB 729 LFGHKPDQDYDLTKAQQLPDKILPDGBETLVSGKSLSSGQARSLAFAVYSRA 788
QY 787 DYLFDPLSLAVDAVHGKHLFENV-IGPKGMKNKRTILVTHSSYLPPOVDVILVMSGGK 845
DB 789 DYLDDILSLAVDAEVSKNILEYVLICKTALKNKKTILLTNVYSILKSGMIALNGE 848
QY 846 ISEMSYQELIAR--DGAFAPFLRTYAS---TEQDQAEANGTVGSGPKAEAYOME 897
DB 849 IVEGNGEDYDWMNRKNTSKLKLLEEDSPIDNGESDVQTEHRSBEVBEF----- 900
QY 898 NGMLVTDSDAKQLOQROUSSSSYSQDISRHNSPAELQKBAKKEFWKL----- 947
DB 901 -----IQLKTESETEDEVLT---ESELRIKNSRRASLATTAPRPFGAQL 945
QY 948 -----MEADKAQTQGVKLVSYYWYDKAIGLFIISFLIFLM-CNHSVALASNYLSLW 999
DB 946 DSVKTAQAKAEKTEVEGRVKTIKYLAIVYKAGV-LGVVLFFLMLTRVPDLAEFWLKYW 1004
QY 1000 TDDPIYNGOEHTRKRLSVYTGALGISQGIAPVGSMAVSI--GGILASRCLHVDLHST 1056
DB 1005 SESENEKSGNERWMMFVGVSILGVAS--AAFNULASIMMLVCSIJSGSKLHESMAKSV 1062
QY 1057 LRSPMSFEERTPGNLVNRFSKELDPTVDSMIPRYIKMFMSLFNVIACVILLATPIA 1116
DB 1063 IRSPMTFEETTPGRILNRSSDMDAVDSMLQYIPSEFPKSLITVLTVALVGNNMPFL 1122
QY 1117 IIPPLGLIYFYQRFVAVSSRQIKRLLESYRSRPVSSHNETLLGVSVIPAFEOEERFIH 1176
DB 1123 VFMNPLVVIYIYQTFPIVLSRELKRLISISYSPIMSLMSESLNGISIIDAYHFERFIY 1182
QY 1177 QSDLKVDENQKAVYPSIVANRMLAVRECVGNCIVLPALPAVI--SRHSLSAGVLGS 1233
DB 1183 LNYEKIQYNDVFEFNEFSTRMRLSVRLQIGATIVATATIALATVMTKQLSSGNGVL 1242
QY 1234 VSYSLQVTTYLTNMLVRRSSEMETNIVAVEELKSESEBEKAPMOIOETAPSSPOVGRV 1293
DB 1243 MSYSLEVTGSLTWIARTVITVETNIVSVERIYVCELPAPQASINPEKRDENKPSKGI 1302
QY 1294 EFRNYSCLRYREDLDVFLRHINVTINGEKVYIGRTGAKSSLTGLFRINISAGAEIIT 1353
DB 1303 EFKVYSKTYRENDPVLNINNVKICEPEKGIYIGRTGAKSTLSALFRILEPTEGKII 1362

QY 1354 DGINIAKIGLHDLRPKITTIIPQDPVLFSGSLRMLNDPFSQYSDEBWTSLFLAHKOFVS 1413
 Db 1363 DGIDISDIDGLFDLRSHLAIIPQDAQAFEGTVKTNLDPFNRYSEDELKRAVEQAHLKPHLE 1422
 QY 1414 AL-----PDKLDHECAEGENLSVGOROLVCLARALLRKTILVLDIA 1456
 Db 1423 KMLHSHKPRGDDSNEDGNVNDILDVKINENGSNLSVGOROLLCLARALLNRSKILVLDIA 1482
 QY 1457 TAAVDLETDDLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRYVLDPKGEIOEYGAPSDILQ 1516
 Db 1483 TASVDMETDKTIQDTIRREFKDRITLITIAHRIDTVLSDKTIIVLDQGSVREFDPSKILS 1542
 QY 1517 OR-GLFYSMADAG 1529
 Db 1543 DKTSIFYSLCEKGG 1556

Search completed: December 15, 2005, 15:36:42
 Job time : 54.8285 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:56:46 ; Search time 243.004 Seconds
(without alignments)
5594.767 Million cell updates/sec

Title: US-10-665-283-6

Perfect score: 9901

Sequence: 1 MALRGCNSADGSDPLMDMNV.....ASVAVAKAKRFSISPSLS 1927

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: uniprot_05.80.*
2: uniprot_sprot.*
3: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7860	79.4	1531	1	MRP1_HUMAN
2	7769	78.5	1515	1	Q9UG99_HUMAN
3	7695	77.7	1531	2	Q864R9_MACTA
4	7689	77.7	1531	2	Q864S0_MACTA
5	7444	75.2	1459	2	Q9UG97_HUMAN
6	7419.5	74.9	1456	2	Q9UQA0_HUMAN
7	7272	73.4	1531	2	Q6UR05_CANPA
8	7242	73.1	1439	2	Q59G19_HUMAN
9	7230.5	73.0	1530	2	Q8HX05_BOVIN
10	7104.5	71.8	1400	2	Q9UG98_HUMAN
11	7002.5	70.7	1528	1	MRP1_MOUSE
12	6932.5	70.0	1532	2	Q810E4_RAT
13	6932.5	70.0	1532	2	Q8CG09_RAT
14	6892	69.6	1523	2	Q810G9_RAT
15	6301	63.6	1525	2	Q5F364_CHICK
16	6075	61.4	1215	2	Q68CP7_HUMAN
17	4526	45.7	1522	2	Q5SUF4_MOUSE
18	4522	45.7	1519	2	Q80ZK8_MOUSE
19	4522	45.7	1522	2	Q59PH0_MOUSE
20	4508.5	45.5	1523	2	Q59DL0_MOUSE
21	4491.5	45.4	1527	1	MRP3_HUMAN
22	4484.5	45.3	1533	2	Q59H05_HUMAN
23	4422	44.7	1514	2	Q96QA9_HUMAN
24	4402	44.5	1498	2	Q59DK9_MOUSE
25	4396	44.4	1522	1	MRP3_RAT
26	4148.5	41.9	1486	2	Q4S7R7_TETNG
27	3885.5	39.2	1548	2	Q7KTC3_DROME
28	3880.5	39.2	1548	2	Q7KTC1_DROME
29	3879.5	39.1	1548	2	Q7KTB7_DROME
30	3874.5	39.1	1548	2	Q7KTC2_DROME
31	3860.5	39.0	1548	2	Q917N0_DROME

32	3857	39.0	1549	2	Q7KTC0_DROME	Q7KTC0_drosophila
33	3845.5	38.8	1548	2	Q8T9C5_DROME	Q8T9C5_drosophila
34	3838.5	38.8	1548	2	Q7KTB8_DROME	Q7KTB8_drosophila
35	3836.5	38.7	1548	2	Q7KTB9_DROME	Q7KTB9_drosophila
36	3818.5	38.6	1548	2	Q7KTD0_DROME	Q7KTD0_drosophila
37	3813.5	38.5	1548	2	Q7KTC8_DROME	Q7KTC8_drosophila
38	3812.5	38.5	1548	2	Q7KTC4_DROME	Q7KTC4_drosophila
39	3807.5	38.5	1548	2	Q7KTC9_DROME	Q7KTC9_drosophila
40	3793.5	38.3	1548	2	Q7KTC7_DROME	Q7KTC7_drosophila
41	3790	38.3	1549	2	Q9VKS6_DROME	Q9VKS6_drosophila
42	3771.5	38.1	1548	2	Q7KTC6_DROME	Q7KTC6_drosophila
43	3769.5	38.1	1548	2	Q8Q998_RAJER	Q8Q998_rajera_erinac
44	3741	37.8	1564	2	Q6PH26_BRAAE	Q6PH26_brachydanio
45	3738.5	37.8	1567	2		

ALIGNMENTS

RESULT 1
MRP1_HUMAN STANDARD; PRT; 1531 AA.
ID MRP1_HUMAN P33527; O14819; P78419;
AC P33527; O14819; P78419;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Multidrug resistance-associated protein 1 (ATP-binding cassette sub-family C member 1).
GN Name=ABCC1; Synonyms=MRP, MRP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE. PubMed=1360704;
RX MEDLINE=93088080; PubMed=1360704;
RA Cole S.P.C., Bhargava G., Gerlach J.H., Mackie J.E., Grant C.E., Almqvist K.C., Stewart A.J., Kurz E.U., Duncan A.M.V., Deeley R.G.;
RT "Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line.";
RL Science 258:1650-1654(1992).
[2]
RP SEQUENCE REVISION.
RX MEDLINE=93262415; PubMed=8098549;
RA Cole S.P.C., Deeley R.G.;
RL "Multidrug resistance-associated protein: sequence correction.";
Science 260:879-879(1993).
[3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96251691; PubMed=8649356;
RA Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P.C., Deeley R.G.;
RL "Structure and expression of the messenger RNA encoding the murine multidrug resistance protein, an ATP-binding cassette transporter.";
Mol. Pharmacol. 49:962-971(1996).
[4]
RP NUCLEOTIDE SEQUENCE, AND ALTERNATIVE SPLICING.
RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1999.5927;
RA Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;
RL "Analysis of the intron-exon organization of the human multidrug-resistance protein gene (MRP) and alternative splicing of its mRNA.";
Genomics 45:368-378(1997).
[5]
RP NUCLEOTIDE SEQUENCE (LARGE SCALE GENOMIC DNA) OF 1131-1531.
RX MEDLINE=9945270; PubMed=10493829; DOI=10.1006/geno.1999.5927;
RA Loftus B.J., Kim U.-J., Shedd V.P., Kalush F., Brandon R., Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L., Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S., Eichler E.B., Harris P.C., Venter J.C., Adams M.D.;
RL "Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q.";
Genomics 60:295-308(1999).

[6] TOPOLOGY, AND CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=9744425; PubMed=9295302; DOI=10.1074/jbc.272.38.23623;
RA Heflinger D.R., Almqvist K.C., Leslie E.M., Gerlach J.H., Grant C.E.,
Deeley R.G., Cole S.P.C.;
RT "Membrane topology of the multidrug resistance protein (MRP). A study
of glycosylation-site mutants reveals an extracytosolic NH2
terminus.";
RL J. Biol. Chem. 272:23623-23630 (1997).
RN [7]
RP TOPOLOGY.
RX MEDLINE=97476249; PubMed=9334225; DOI=10.1074/jbc.272.42.26479;
RA Kaet C., Gros P.;
RT "Topology mapping of the amino-terminal half of multidrug resistance-
associated protein by epitope insertion and immunofluorescence.";
RL J. Biol. Chem. 272:26479-26487 (1997).
RN [8]
RP TOPOLOGY.
RX MEDLINE=98153110; PubMed=9485377; DOI=10.1021/bi972332v;
RA Kaet C., Gros P.;
RT "Epitope insertion favors a six transmembrane domain model for the
carboxy-terminal portion of the multidrug resistance-associated
protein.";
RT Biochemistry 37:2305-2313 (1998).
RN [9]
RP MUTAGENESIS OF ASP-792; ASP-793; LYS-1333 AND 1454-ASP-1455.
RX MEDLINE=21362977; PubMed=11469806; DOI=10.1006/abbi.2001.2441;
RA Cui L., Hou Y.-X., Riordan J.R., Chang X.-B.;
RT "Mutations of the Walker B motif in the first nucleotide binding
domain of multidrug resistance protein MRP1 prevent conformational
maturation.";
RL Arch. Biochem. Biophys. 392:153-161 (2001).
RN [10]
RP MUTAGENESIS OF TRP-1246.
RX MEDLINE=21238219; PubMed=11278867; DOI=10.1074/jbc.M011246200;
RA Ito K., Olsen S.L., Qiu W., Deeley R.G., Cole S.P.C.;
RT "Mutation of a single conserved tryptophan in multidrug resistance
protein 1 (MRP1/ABCC1) results in loss of drug resistance and
selective loss of organic anion transport.";
RL J. Biol. Chem. 276:15616-15624 (2001).
RN [11]
RP VARIANTS GLN-633 AND VAL-671.
RX MEDLINE=20296630; PubMed=10835642; DOI=10.1038/76102;
RA Le Saux O., Urban Z., Tschuch C., Csiszar K., Bacchelli B.,
Quaglini D., Pasquali-Ronchetti I., Pope F.M., Richards A., Terry S.,
Bercovich L., de Paeppe A., Boyd C.D.;
RT "Mutations in a gene encoding an ABC transporter cause pseudoxanthoma
elasticum.";
RL Nat. Genet. 25:223-227 (2000).
RN [12]
RP VARIANT VAL-671.
RX MEDLINE=20283940; PubMed=10811882; DOI=10.1073/pnas.100041297;
RA Ringpfeil F., Lebowitz M.G., Christiano A.M., Uitto J.;
RT "Pseudoxanthoma elasticum: mutations in the MRP6 gene encoding a
transmembrane ATP-binding cassette (ABC) transporter.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:6001-6006 (2000).
RN [13]
RP VARIANT SER-433, AND CHARACTERIZATION OF VARIANT VAL-671.
RX MEDLINE=21578494; PubMed=11721885; DOI=10.1007/s100380170017;
RA Conrad S., Kaufmann H.-M., Ito K., Deeley R.G., Cole S.P.C.,
Schrenk D.;
RT "Identification of human multidrug resistance protein 1 (MRP1)
mutations and characterization of a G671V substitution.";
RL J. Hum. Genet. 46:656-663 (2001).
RN [14]
RP VARIANTS MET-117 AND LEU-1512.
RX DOI=10.1002/1098-1004(2000)17:1<74>:AID-HUMU14>3.0.CO;2-F;
RA Perdu J., Germain D.P.;
RT "Identification of novel polymorphisms in the PMS and MRP1 (ABCC1)
genes at locus 16p13.1 and exclusion of both genes as responsible for
pseudoxanthoma elasticum.";
RL Hum. Mutat. 17:74-75 (2001).

[15]
RP VARIANTS SER-433; ILE-73; GLN-723 AND GLN-1058.
RX MEDLINE=21163848; PubMed=11266082;
RA DOI=10.1097/00008571-200103000-00008;
RX Ito S., Ito I., Tanabe M., Suzuki A., Higuchi S., Otsubo K.;
RT "Polymorphism of the ABC transporter genes, MRP1, MRP2 and MRP2/CMOAT,
in healthy Japanese subjects.";
RL Pharmacogenetics 11:175-184 (2001).
CC -1- FUNCTION: May participate directly in the active transport of
drugs into subcellular organelles or influence drug distribution
indirectly.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=8;
CC Comment-Additional isoforms seem to exist. Experimental
CC confirmation may be lacking for some isoforms;
CC Name=Alllexons;
CC IsoId=P33527-1; Sequence=Displayed;
CC Name=Delexon-17;
CC IsoId=P33527-2; Sequence=VSP_000037;
CC Name=Delexon-18;
CC IsoId=P33527-3; Sequence=VSP_000038;
CC Name=Delexon-30;
CC IsoId=P33527-4; Sequence=VSP_000039;
CC Name=Delexon-17-18;
CC IsoId=P33527-5; Sequence=VSP_000037, VSP_000038;
CC Name=Delexon-17-30;
CC IsoId=P33527-6; Sequence=VSP_000037, VSP_000039;
CC Name=Delexon-18-30;
CC IsoId=P33527-7; Sequence=VSP_000038, VSP_000039;
CC Name=Delexon-17-18-30;
CC IsoId=P33527-8; Sequence=VSP_000037, VSP_000039;
CC -1- TISSUE SPECIFICITY: Lung, testis and peripheral blood mononuclear
cells.
CC -1- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.
CC -1- SIMILARITY: Contains 2 ABC transmembrane type-1 domains.
CC -1- SIMILARITY: Contains 2 ABC transporter domains.
CC -1- DATABASIS: NME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobogen.fr/services/chromocancer/genes/MRPID106.html".
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL, L05628; AAB4616.1; -, mRNA.
CC EMBL, AF022853; AAB83983.1; -, Genomic DNA.
CC EMBL, AF022824; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022825; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022826; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022827; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022828; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022829; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022830; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022831; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022832; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022833; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022834; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022835; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022836; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022837; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022838; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022839; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022840; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022841; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022842; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022843; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022844; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022845; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022846; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022847; AAB83983.1; JOINED; Genomic DNA.
CC EMBL, AF022848; AAB83983.1; JOINED; Genomic DNA.

DR EMBL; AF022849; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022850; AAB83983.1; JOINED; Genomic DNA.

Query Match 79.4%; Score 7860; DB 1; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MALRGFCSDADSPDLMNMTNMTNNDPFCFQNTLVVWPCGYLACEPFYLSRH 60
DB 1 MALRGFCSDADSPDLMNMTNMTNNDPFCFQNTLVVWPCGYLACEPFYLSRH 60
QY 61 DRGYIQMTPLNKTKTALGFLIMTYCMADLFYSFWEBSRGIFLAPVFLVSTLIGITLLA 120
DB 61 DRGYIQMTPLNKTKTALGFLIMTYCMADLFYSFWEBSRGIFLAPVFLVSTLIGITLLA 120
QY 121 TELIOERRKGVSSGIMLTFWLVALCALAIRSKITMTAKEDQVDFRDTFFYFSS 180
DB 121 TELIOERRKGVSSGIMLTFWLVALCALAIRSKITMTAKEDQVDFRDTFFYFSS 180
QY 181 LLLIOVLSCFSDRSPFSETIHDNPPCSSASFLSRITFWMTGLIYRGYQPLEGSD 240
DB 181 LLLIOVLSCFSDRSPFSETIHDNPPCSSASFLSRITFWMTGLIYRGYQPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNWKKECAKTRKQPVYVSSKDPQPKSSKYDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNWKKECAKTRKQPVYVSSKDPQPKSSKYDANEVEAL 300
QY 301 IVKSPKEMNPSPFLKYLTGFGYFLMSFFPKAIDHLMFSGPOILKLIKYNDYKAP 360
DB 301 IVKSPKEMNPSPFLKYLTGFGYFLMSFFPKAIDHLMFSGPOILKLIKYNDYKAP 360
QY 361 WQGFYTVLLFVTAQOTLVHOFPHICFVSGMRKTAIVGAYRKALVITNSARKSTV 420
DB 361 WQGFYTVLLFVTAQOTLVHOFPHICFVSGMRKTAIVGAYRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOVILALYLLMTLGPVSLAGVAVMLMPVN 480
DB 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOVILALYLLMTLGPVSLAGVAVMLMPVN 480
QY 481 AVAMAKTKTYQVAHMSKDNRIKLMNEILINGIKLVKLYAMELAFKQVLAIRQBELVK 540
DB 481 AVAMAKTKTYQVAHMSKDNRIKLMNEILINGIKLVKLYAMELAFKQVLAIRQBELVK 540
QY 541 KSAVLSAVGTFVTCPELVALCTPAVYVITIDENIILDAQTAVSALFNIILRPPLNIP 600
DB 541 KSAVLSAVGTFVTCPELVALCTPAVYVITIDENIILDAQTAVSALFNIILRPPLNIP 600
QY 601 MVTSSIVQASVSLKRLRIFLSHEELBPSIERRPVKDGGSNTSIYNNATFTWARSDDPT 660
DB 601 MVTSSIVQASVSLKRLRIFLSHEELBPSIERRPVKDGGSNTSIYNNATFTWARSDDPT 660
QY 661 LNSITFSIPEGALVAVVGQVCGKSSILSALLAEMDKVEGHVAIKGSVAVYPOAMIOND 720
DB 661 LNSITFSIPEGALVAVVGQVCGKSSILSALLAEMDKVEGHVAIKGSVAVYPOAMIOND 720
QY 721 SLRENLILFGQLEPPYRSYIQAALLPDLIELIPSGRTIEGKCVNLGGQKQORSILAR 780
DB 721 SLRENLILFGQLEPPYRSYIQAALLPDLIELIPSGRTIEGKCVNLGGQKQORSILAR 780
QY 781 AAVSNADIVLFDPLSLAVDAVHGKHIPENYIGPKMKDKNTRILVTHSMGYLROVDYIY 840
DB 781 AAVSNADIVLFDPLSLAVDAVHGKHIPENYIGPKMKDKNTRILVTHSMGYLROVDYIY 840
QY 841 MSGGKISEMGSYOELLARDGAPAFPLRTYASTBOEDAEENGTVGSGPGEAKQOMENG 900
DB 841 MSGGKISEMGSYOELLARDGAPAFPLRTYASTBOEDAEENGTVGSGPGEAKQOMENG 900
QY 901 LVTHDSAGKQOROLSSSSSYSGDISRHNNSTALQRAKAKERTWKLMDKXQQTQGVKL 960
DB 901 LVTHDSAGKQOROLSSSSSYSGDISRHNNSTALQRAKAKERTWKLMDKXQQTQGVKL 960
QY 961 SVYWDYKAIKGLFISFLIFLFCMCHVSALASNYMLSLMTDDEIVNGTOHTKVRLSVYG 1020
DB 961 SVYWDYKAIKGLFISFLIFLFCMCHVSALASNYMLSLMTDDEIVNGTOHTKVRLSVYG 1020

```

```

DB 961 SVYWDYKAIKGLFISFLIFLFCMCHVSALASNYMLSLMTDDEIVNGTOHTKVRLSVYG 1020
QY 1021 ALGISGIANVFGSMVSTGILLASRCIHYDLHSLIRSPMSFPEPTPSGNLVNREKEL 1080
DB 1021 ALGISGIANVFGSMVSTGILLASRCIHYDLHSLIRSPMSFPEPTPSGNLVNREKEL 1080
QY 1081 DTVDMSIPEVIYKMFMSLNFVIGACIVILLATPIAIIIPPLGLIYFVQRFVASSROL 1140
DB 1081 DTVDMSIPEVIYKMFMSLNFVIGACIVILLATPIAIIIPPLGLIYFVQRFVASSROL 1140
QY 1141 KRLESVSRSPVSHFNETLGVSVIRAFEBQERFIHQSDLKVDENQKAYVPSIVANRWLA 1200
DB 1141 KRLESVSRSPVSHFNETLGVSVIRAFEBQERFIHQSDLKVDENQKAYVPSIVANRWLA 1200
QY 1201 VRLCEVNCIVLPAALPAVYSRSLAGLVGVSISLOVTTYLNLVYKSSSEMETRYVA 1260
DB 1201 VRLCEVNCIVLPAALPAVYSRSLAGLVGVSISLOVTTYLNLVYKSSSEMETRYVA 1260
QY 1261 VERLKEYSETEKAPMOIOETAPSSMPQVGRVFPFNVCILRYREDLDFVLRHINVTINGG 1320
DB 1261 VERLKEYSETEKAPMOIOETAPSSMPQVGRVFPFNVCILRYREDLDFVLRHINVTINGG 1320
QY 1321 EKVGIYGRGAGKSSVTLGLFRINESAGEIIIDGINIAKIGHDLRFKIIIPQPVLF 1380
DB 1321 EKVGIYGRGAGKSSVTLGLFRINESAGEIIIDGINIAKIGHDLRFKIIIPQPVLF 1380
QY 1381 SCSLRNMLDPSQYSDPEBWTSLLELAHKDFVSALPDKLDHBCAEGENISVQOROLVCL 1440
DB 1381 SCSLRNMLDPSQYSDPEBWTSLLELAHKDFVSALPDKLDHBCAEGENISVQOROLVCL 1440
QY 1441 ARALLKTKTILVDEAPTAADLETDDLIOSTRTOPEDECTVLTARLNTIMTYRYIVL 1500
DB 1441 ARALLKTKTILVDEAPTAADLETDDLIOSTRTOPEDECTVLTARLNTIMTYRYIVL 1500
QY 1501 DKGEIOEYGA PSDLLQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGA PSDLLQORGLFYSMAKDAGLV 1531

RESULT 2
Q9U099 HUMAN
ID Q9U099 HUMAN PRELIMINARY; PRT; 1515 AA.
AC Q9U099
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Multidrug resistance protein (Fragment).
GN Name=MRP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RS NUCLEOTIDE SEQUENCE.
RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;
RA Grant C.B., Kurz E.U., Cole S.P.C., Deley R.G.;
RT "Analysis of the intron-exon organization of the human multidrug-
   resistance protein gene (MRP) and alternative splicing of its mRNA.";
RL Genomics 45:368-378(1997).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DB EMBL; AF022853; AAB83979.1; -, Genomic DNA.
DB EMBL; AF022827; AAB83979.1; JOINED; Genomic DNA.
DB EMBL; AF022828; AAB83979.1; JOINED; Genomic DNA.
DB EMBL; AF022829; AAB83979.1; JOINED; Genomic DNA.
DB EMBL; AF022831; AAB83979.1; JOINED; Genomic DNA.
DB EMBL; AF022833; AAB83979.1; JOINED; Genomic DNA.
DB EMBL; AF022835; AAB83979.1; JOINED; Genomic DNA.
DB EMBL; AF022837; AAB83979.1; JOINED; Genomic DNA.
DB EMBL; AF022839; AAB83979.1; JOINED; Genomic DNA.
DB EMBL; AF022841; AAB83979.1; JOINED; Genomic DNA.
DB EMBL; AF022850; AAB83979.1; JOINED; Genomic DNA.
DB EMBL; AF022849; AAB83979.1; JOINED; Genomic DNA.

```

DR EMBL; AF022848; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022847; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022846; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022845; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022844; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022843; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022842; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022841; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022840; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022839; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022838; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022837; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022836; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022835; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022834; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022833; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022832; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022831; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022830; AAB83979.1; JOINED; Genomic DNA.
DR HSP; F08716; 1MT0.
DR Ensembl; ENSG00000103222; Homo sapiens.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000186; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC membrane 1.
DR InterPro; IPR001140; ABC TM transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; NRP_assoc.
DR InterPro; IPR00719; Prot_kinase.
DR Pfam; PF00664; ABC_tran; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR Prodom; PD000006; ABC transporter; 2.
DR SMART; SM00382; AAA_2.
DR TIGRFAMs; TIGR00957; NRP_assoc_pro; 1.
DR PROSITE; PS00929; ABC_TM1F; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN KINASE_ATP; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding.
FT NON_TER 1 1
SQ SEQUENCE 1515 AA; 169853 MW; DC95528217C439FE CRC64;
Query Match 78.5%; Score 7769; DB 2; Length 1515;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 377 QTVLHGYFHI CRYSGMRITANVIGAVYRKALVITNSARKSSVYGEIVNLSVDQGRFMD 436
DB 361 QTVLHGYFHI CRYSGMRITANVIGAVYRKALVITNSARKSSVYGEIVNLSVDQGRFMD 420
QY 437 LATYINWISAPLOVIALYLLMLNIGPSYLAGAVVWVLPVNAVAMKTKTYQVAHMK 496
DB 421 LATYINWISAPLOVIALYLLMLNIGPSYLAGAVVWVLPVNAVAMKTKTYQVAHMK 480
QY 497 SKNRIKLAMEIINGIKVLKLYAMELAFKDKVLAIROBELKVLKKSAYLSAVGTFTWCT 556
DB 481 SKNRIKLAMEIINGIKVLKLYAMELAFKDKVLAIROBELKVLKKSAYLSAVGTFTWCT 540
QY 557 PELVALCTPVAVYTTIDENNILDAQTAFAVSLAFNIRLPFNILPMVYSSIVQASVSLKRL 616
DB 541 PELVALCTPVAVYTTIDENNILDAQTAFAVSLAFNIRLPFNILPMVYSSIVQASVSLKRL 600
QY 617 RIFLSHELEPDSIERPVDGGGNTSITVRNATFTMARSDPPLNIGITSIPGALVAV 676
DB 601 RIFLSHELEPDSIERPVDGGGNTSITVRNATFTMARSDPPLNIGITSIPGALVAV 660
QY 677 VGVGCGKSSILSALTAEMLKVEGHVAIKGSVAVVPOQAMIQNDLSRENILFGQLEBEPY 736
DB 661 VGVGCGKSSILSALTAEMLKVEGHVAIKGSVAVVPOQAMIQNDLSRENILFGQLEBEPY 720
QY 737 YRSVYQACALLPDLIELIPSGDRTEIGKGVNLSGGQKQVSLARAVYSNADYILFDPLS 796
DB 721 YRSVYQACALLPDLIELIPSGDRTEIGKGVNLSGGQKQVSLARAVYSNADYILFDPLS 780
QY 797 AVNAHVGHKLFENVIGPKGMLKXKTRILYTHSNYSYLPQVAVIIVMSGGKISEMGSYOEL 856
DB 781 AVNAHVGHKLFENVIGPKGMLKXKTRILYTHSNYSYLPQVAVIIVMSGGKISEMGSYOEL 840
QY 857 ARDGAFAEPLRTYASTEOBDAEENGTVGSGPGEAKQOMENGLVYDSAGKOLORLSS 916
DB 841 ARDGAFAEPLRTYASTEOBDAEENGTVGSGPGEAKQOMENGLVYDSAGKOLORLSS 900
QY 917 SSSYSGDISHHNSTAELOQAERKERTWKLMEADKAQTCQVKLSVWMDYKALGLRISF 976
DB 901 SSSYSGDISHHNSTAELOQAERKERTWKLMEADKAQTCQVKLSVWMDYKALGLRISF 960
QY 977 LSTFLPCHNVASLASVWYLSLMTDPIVNGTQHTVRSLSVYGALGISGIAVFGSMA 1036
DB 961 LSTFLPCHNVASLASVWYLSLMTDPIVNGTQHTVRSLSVYGALGISGIAVFGSMA 1020
QY 1037 VSIGILASRCLHVDLHSLRSFMSFEFRTPSGNLVNRSKELDYDSMIPVYIKAFMG 1096
DB 1021 VSIGILASRCLHVDLHSLRSFMSFEFRTPSGNLVNRSKELDYDSMIPVYIKAFMG 1080
QY 1097 SLENVIGACTIVIIILATPIAIIIPPLGLIYFPVQRFYVASSROLKRLSVSRSPVSHFN 1156
DB 1081 SLENVIGACTIVIIILATPIAIIIPPLGLIYFPVQRFYVASSROLKRLSVSRSPVSHFN 1140
QY 1157 ETLIGSVIRAFEEQEFTHQSDLKVDENOKAYYPSIVARWMLAVRLECVCNCLVPAAL 1216
DB 1141 ETLIGSVIRAFEEQEFTHQSDLKVDENOKAYYPSIVARWMLAVRLECVCNCLVPAAL 1200
QY 1217 FAVISRSLSAGLVGLSVSLQVTTYLNLVNRSSMEENTNIVABERLKESETEKEAPW 1276
DB 1201 FAVISRSLSAGLVGLSVSLQVTTYLNLVNRSSMEENTNIVABERLKESETEKEAPW 1260
QY 1277 QIETAPSSWPQVGRVPEFRNYCLARYREDLPVLARHINVTINGEKVIGVGRTAGKSSL 1336
DB 1261 QIETAPSSWPQVGRVPEFRNYCLARYREDLPVLARHINVTINGEKVIGVGRTAGKSSL 1320
QY 1337 TLGLFRINESAGEIITIDGINIAKIGLHDLRFKTTIIPQPVLPFGSLRNMLDPFSQYSD 1396
DB 1321 TLGLFRINESAGEIITIDGINIAKIGLHDLRFKTTIIPQPVLPFGSLRNMLDPFSQYSD 1380
QY 1397 BEVWTSLELHLKDFVSLPDKLDHECABEGENLSVGOQVCLARALRLKTKLIVLDEA 1456
DB 1381 BEVWTSLELHLKDFVSLPDKLDHECABEGENLSVGOQVCLARALRLKTKLIVLDEA 1440
QY 1457 TAAVDETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQ 1516

Db 1441 TAAVDLEDDLIQSTIRTFQEDCTVLTFAHRLNTIMDYTRIVYLDKSGIQEYGAPSDLLQ 1500
 QY 1517 ORGLFYMAKXAGLV 1531
 Db 1501 ORGLFYMAKXAGLV 1515

RESULT 3
 08649 MACFA
 ID 08649 MACFA PRELIMINARY; PRT; 1531 AA.
 AC 08649;
 DT 01-JUN-2003 (TREMBlurel. 24, Created)
 DT 01-JUN-2003 (TREMBlurel. 24, Last sequence update)
 DT 01-MAR-2004 (TREMBlurel. 26, Last annotation update)
 DE Multidrug resistance protein 1B.
 GN Name=MRP1;
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 OC Cercopithecoidea; Cercopithecinae; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP NOCLEOTIDE SEQUENCE.
 RX MEDLINE=22544876; PubMed=12657726;
 RA Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,
 RA Dantzig A.H., Perry W.L.,
 RT "Cloning and functional characterization of the multidrug resistance-
 RT associated protein (MRP1/ABCC1) from the cynomolgus monkey.";
 RL Mol. Cancer Ther. 2:307-316(2003).
 RN [2]
 RP NOCLEOTIDE SEQUENCE.
 RU Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- Similarity: Belongs to the ABC transporter family.
 DR EMBL; AY146673; AAN65349.1; -; mRNA.
 DR HSSP; P08716; IMTO.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0016887; F:ATPase activity; IEA.
 DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
 DR GO; GO:0000166; F:nucleotide binding; IEA.
 DR GO; GO:0006810; P:transporter; IEA.
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR011527; ABC_membrane_1.
 DR InterPro; IPR001140; ABC_TM_transp.
 DR InterPro; IPR003439; ABC_transp_like.
 DR InterPro; IPR005292; MRP_assoc.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00664; ABC_membrane; 2.
 DR Pfam; PF00005; ABC_tran; 2.
 DR ProDom; PD000006; ABC_transporter; 2.
 DR SMART; SM00382; AAA; 2.
 DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
 DR PROSITE; PSS0929; ABC_TM1F; 2.
 DR PROSITE; PSS0211; ABC_TRANSPORTER_1; 2.
 DR PROSITE; PSS0893; ABC_TRANSPORTER_2; 2.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR KEGG; K00001; PROTEIN_KINASE_ATP; UNKNOWN_1.
 KW ATP-binding; Nucleotide-binding; Repeat.
 SQ SEQUENCE 1531 AA; 171659 MW; 1AE788E9DF9EF459 CRC64;

Query Match 77.7%; Score 7695; DB 2; Length 1531;
 Best Local Similarity 97.5%; Pred. No. 0;
 Matches 1493; Conservative 20; Mismatches 18; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTVNTSNDPTKCFONTVTVVWPCFYLMACFPFYFLYSRH 60
 Db 1 MALRGFCSADGSDPLMDMNTVNTSNDPTKCFONTVTVVWPCFYLMACFPFYFLYSRH 60
 QY 61 DRGIVQMTPLNKTKTALIGLLIVCWADLFYSFWEISRGIFLAPVPLVSPTLLGITTLA 120
 Db 61 DRGIVQMTPLNKTKTALIGLLIVCWADLFYSFWEISRGIFLAPVPLVSPTLLGITTLA 120

QY 121 TFLIQLERRKGVSSGIMLFWMVAALCALAIILRSKIMTALKEDAQVDLFRDITFEYYFS 180
 Db 121 TFLIQLERRKGVSSGIMLFWMVAALCALAIILRSKIMTALKEDAQVDLFRDITFEYYFS 180
 QY 181 LLLIQLVLSGFSRDRPLPSETIHDNPNCPSSASFLSRITFWMTGLIVRGYRPLRSGD 240
 Db 181 LLLIQLVLSGFSRDRPLPSETIHDNPNCPSSASFLSRITFWMTGLIVRGYRPLRSGD 240
 QY 241 LMSLNKEDTSEQVVPVTVKMKKCATRQKQPVVVVYSSKDPAPKSSSKVDANEVEAL 300
 Db 241 LMSLNKEDTSEQVVPVTVKMKKCATRQKQPVVVVYSSKDPAPKSSSKVDANEVEAL 300
 QY 301 IVKSPQKEMNPSPKLVYKTFGPFLMSFPFKIHDMFSGPOILLIKLIFVNDTAPD 360
 Db 301 IVKSPQKEMNPSPKLVYKTFGPFLMSFPFKIHDMFSGPOILLIKLIFVNDTAPD 360
 QY 361 MOGYFTVLLFVTAQLOTLVLAHQYFHCIPVSGRIKTAIVGAVERKALVITNSARKSTV 420
 Db 361 MOGYFTVLLFVTAQLOTLVLAHQYFHCIPVSGRIKTAIVGAVERKALVITNSARKSTV 420
 QY 421 GEIVNLSVDAQRPMDLATTYNNIMWAPLOVILALYLLMNLGFSVLGAVAVMLVPVN 480
 Db 421 GEIVNLSVDAQRPMDLATTYNNIMWAPLOVILALYLLMNLGFSVLGAVAVMLVPVN 480
 QY 481 AVNAMKTKTYQVAHMSKDKRIKLMNLIIGIKVLKLYAMELAFKDKVALROBELVYLK 540
 Db 481 AVNAMKTKTYQVAHMSKDKRIKLMNLIIGIKVLKLYAMELAFKDKVALROBELVYLK 540
 QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITIDENNILDAOTAFVSLAFNILRPLNLP 600
 Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITIDENNILDAOTAFVSLAFNILRPLNLP 600
 QY 601 MVISSIVQASVSLKRLRIFLSHEBLEPDSIERRPVKDGGSGTNSITVNNATFTMARSDPT 660
 Db 601 MVISSIVQASVSLKRLRIFLSHEBLEPDSIERRPVKDGGSGTNSITVNNATFTMARSDPT 660
 QY 661 LMGITTSIPRGALVAVVGQVCGKSSLSLALBMDVBEHVAKSGVAVVPQAMVOND 720
 Db 661 LMGITTSIPRGALVAVVGQVCGKSSLSLALBMDVBEHVAKSGVAVVPQAMVOND 720
 QY 721 SLRENILFGCLPEEPYRSVIOACALLPDLIELIPSGRTEIGEGVNLSGGQKORVSLAR 780
 Db 721 SLRENILFGCLPEEPYRSVIOACALLPDLIELIPSGRTEIGEGVNLSGGQKORVSLAR 780
 QY 781 AVYSNADITYLPDPLSADVAHVGKHIPEVNTIGPKMKLKNKTRILVTHSMYSLPQVDVIY 840
 Db 781 AVYSNADITYLPDPLSADVAHVGKHIPEVNTIGPKMKLKNKTRILVTHSMYSLPQVDVIY 840
 QY 841 MSGGKISEMGSYOELLARDGAFAPLRTYVASTROBODAEENGVTGVSGPKKAKOMENGM 900
 Db 841 MSGGKISEMGSYOELLARDGAFAPLRTYVASTROBODAEENGVTGVSGPKKAKOMENGM 900
 QY 901 LVTDSAGKQLOROLSSSSSYSGDISRHHNSTAELOKAEKEFTKMLEADKATOGVYL 960
 Db 901 LVTDSAGKQLOROLSSSSSYSGDISRHHNSTAELOKAEKEFTKMLEADKATOGVYL 960
 QY 961 SVYWDYMKALGIFISPLIFLFCNCHVSAALSNYMLSLMTDDBIVNGTOGHTVRLSVYG 1020
 Db 961 SVYWDYMKALGIFISPLIFLFCNCHVSAALSNYMLSLMTDDBIVNGTOGHTVRLSVYG 1020
 QY 1021 ALGISGIAVFGSMASVSTGILLASRCIHVDLHSLIRSMSPFEKTPSGNLVNRPSKEL 1080
 Db 1021 ALGISGIAVFGSMASVSTGILLASRCIHVDLHSLIRSMSPFEKTPSGNLVNRPSKEL 1080
 QY 1081 DTVDSMIPEYIKPMFMSLFVIGACIVILLATPIAIIIPPLGLIYEFVORFYVASSROL 1140
 Db 1081 DTVDSMIPEYIKPMFMSLFVIGACIVILLATPIAIIIPPLGLIYEFVORFYVASSROL 1140
 QY 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENOKAYYPISVANRWLA 1200
 Db 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENOKAYYPISVANRWLA 1200
 QY 1201 VRLCEVGNCTVLPALPAVVISRHSLSAGLVGSLVSLSQVTTYTLNMLVNMSSMEETNIYA 1260

```

Db      1201 VRLECVANCIVLFAFALPAVLSRHSLSAGLVGLSVSYSIQVTTYTNMVLWRSMSEMETIVA 1260
Qy      1261 VERLKESETEKEAPMOIOETAPSPSWPQVGRVVERPNCLAREDLDPVLRHINVTINGG 1320
Db      1261 VERLKESETEKEAPMOIOETAPSPSWPQVGRVVERPNCLAREDLDPVLRHINVTINGG 1320
Qy      1321 EKVGVIRGTGAGKSSLTGLFRINESAGEIIGININAKTGLDHLRDKITIIIPQDPVLF 1380
Db      1321 EKVGVIRGTGAGKSSLTGLFRINESAGEIIGININAKTGLDHLRDKITIIIPQDPVLF 1380
Qy      1381 SGSRPMNLDPSSQYSDEERWMTSLSLAHKDFVSALPDKLDHECAEGENLSVGQROLVCL 1440
Db      1381 SGSRPMNLDPSSQYSDEERWMTSLSLAHKDFVSALPDKLDHECAEGENLSVGQROLVCL 1440
Qy      1441 ARALLRKTKIIVLBATAVAVLETDLDLQSTIRIQFEDCTVLTIAHRLNTIMDYTRVYL 1500
Db      1441 ARALLRKTKIIVLBATAVAVLETDLDLQSTIRIQFEDCTVLTIAHRLNTIMDYTRVYL 1500
Qy      1501 DKGEIOEGAPSDLLQORGLFYMAKQAGLV 1531
Db      1501 DKGEIOEGAPSDLLQORGLFYMAKQAGLV 1531

```

RESULT 4

Q864S0_MACFA PRELIMINARY; PRT; 1531 AA.

```

ID      0864S0_MACFA PRELIMINARY; PRT; 1531 AA.
AC      0864S0:
DT      01-JUN-2003 (TREMblrel. 24, Created)
DT      01-JUN-2003 (TREMblrel. 24, Last sequence update)
DT      01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE      Multidrug resistance protein 1A.
OS      Name=MRP1;
OS      Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC      Cercopithecoidea; Cercopithecinae; Macaca.
OX      NCBI_TaxID=9541;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      TISSUE=Kidney;
RX      MEDLINE=22544876; PubMed=12657726;
RA      Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,
RA      Danczig A.H., Perry W.L.;
RT      "Cloning and functional characterization of the multidrug resistance-
RT      associated protein (MRP1/ABCC1) from the cynomolgus monkey.";
RL      Mol. Cancer Ther. 2:307-316(2003).
RN      [2]
RP      NUCLEOTIDE SEQUENCE.
RC      TISSUE=Kidney;
RA      Perry W.L., Ili, Godinot N.;
RL      Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC      -!- SIMILARITY: Belongs to the ABC transporter family.
DR      EMBL; AY146672; AAN65348.1; -; mRNA.
DR      HSSP; P08716; IMT0.
DR      GO; GO:0016021; C:integral to membrane; IEA.
DR      GO; GO:0005524; F:ATP binding; IEA.
DR      GO; GO:0016887; F:ATPase activity; IEA.
DR      GO; GO:0042626; F:ATPase activity; coupled to transmembrane m. . .; IEA.
DR      GO; GO:0000166; F:nucleotide binding; IEA.
DR      InterPro; IPR003593; AAA_ATPase.
DR      InterPro; IPR011527; ABC_membrane_1.
DR      InterPro; IPR001140; ABC_TM_transpDc.
DR      InterPro; IPR003439; ABC_transp_like.
DR      InterPro; IPR005292; MRP_assoc.
DR      Pfam; PF00664; ABC_membrane; 2.
DR      Pfam; PF00005; ABC_tran; 2.
DR      ProDom; PD000066; ABC_transporter; 2.
DR      SMART; SM00382; AAA; 2.
DR      TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR      ProSite; PS50929; ABC_TMI; 2.

```

```

DR      PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR      PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR      PROSITE; PS00107; PROTEIN KINASE_ATP; UNKNOWN_1.
KW      ATP-binding; Nucleotide-binding; Repeat.
SQ      SEQUENCE 1531 AA; 171701 MW; 168712B8CC2D2B89 CRC64;

```

Query Match 77.7%; Score 7689; DB 2; Length 1531;
 Best local similarity 97.6%; Pred. No. 0;
 Matches 1494; Conservative 18; Mismatches 19; Indels 0; Gaps 0;

```

Qy      1 MALRGPSADGSDPLMDMNTMTNSNDPTKCPONTLVVWPCCYVLMACPEFYLYSRH 60
Db      1 MALRGPSADGSDPLMDMNTMTNSNDPTKCPONTLVVWPCCYVLMACPEFYLYSRH 60
Qy      61 DRGYQMTPLNKTATLGFLLMIVCMADLFYSFMERSRGIFLAPVFLVSPFLIGITLLTA 120
Db      61 DRGYQMTPLNKTATLGFLLMIVCMADLFYSFMERSRGIFLAPVFLVSPFLIGITLLTA 120
Qy      121 TFLIQLEERRKGVQSSGIMLTFWLVAVCALALIRSKIMTALKEDAVQDLFRDITFYVFS 180
Db      121 TFLIQLEERRKGVQSSGIMLTFWLVAVCALALIRSKIMTALKEDAVQDLFRDITFYVFS 180
Qy      181 LLLIQVLSCGSDSPLEFSETHDNPCCPSSASFLSRITFMWTTGLIVRGYRQPLEGSD 240
Db      181 LLLIQVLSCGSDSPLEFSETHDNPCCPSSASFLSRITFMWTTGLIVRGYRQPLEGSD 240
Qy      241 LMSLNKEDTSEQVVPVLYKMKKSCAKTRKQPVKVVYSSKDPAQPKSSKVDANEVEAL 300
Db      241 LMSLNKEDTSEQVVPVLYKMKKSCAKTRKQPVKVVYSSKDPAQPKSSKVDANEVEAL 300
Qy      301 IVKSPQKEMNPSPLEKVLTKTFGPFLMSFFPKAIDHLMFSGPOLIKLIFVNDTYAPD 360
Db      301 IVKSPQKEMNPSPLEKVLTKTFGPFLMSFFPKAIDHLMFSGPOLIKLIFVNDTYAPD 360
Qy      361 WQGFYFVLLFVACLOTVLVHQYFHIQVSGMRKIKAVIGAVYRKALVTNARKSTV 420
Db      361 WQGFYFVLLFVACLOTVLVHQYFHIQVSGMRKIKAVIGAVYRKALVTNARKSTV 420
Qy      421 GEIVNLSVDAQREMDLATYINMIMSAPLOVILLALYLMLNTGSPVLAVAVMVLMPVN 480
Db      421 GEIVNLSVDAQREMDLATYINMIMSAPLOVILLALYLMLNTGSPVLAVAVMVLMPVN 480
Qy      481 AVMAKTKTYQVAHMKSKDNRIKLMNSILINGIKVLKYAMELAFKDYALAIROELKVLK 540
Db      481 AVMAKTKTYQVAHMKSKDNRIKLMNSILINGIKVLKYAMELAFKDYALAIROELKVLK 540
Qy      541 KSAVLSAVGFTTWCTPFVVALCTFAVYVITDENIIDAQTA PSLALPNTLRPPLILP 600
Db      541 KSAVLSAVGFTTWCTPFVVALCTFAVYVITDENIIDAQTA PSLALPNTLRPPLILP 600
Qy      601 MVISIYQASVSLKRLIFLSHEELBPSIERRPVKDGGGNSITVENATFTMARSDPPT 660
Db      601 MVISIYQASVSLKRLIFLSHEELBPSIERRPVKDGGGNSITVENATFTMARSDPPT 660
Qy      661 LMGITFSIPREGALVAVVGVQCGKSSLLSALLAEMDKVEGHVALKGSVAVYPQQAMTQND 720
Db      661 LMGITFSIPREGALVAVVGVQCGKSSLLSALLAEMDKVEGHVALKGSVAVYPQQAMTQND 720
Qy      721 SLRNNILPFCGLEPPYRSVYQAACALLPDLEILPSSGRTIEGEGVNLSSGQKRVSLAR 780
Db      721 SLRNNILPFCGLEPPYRSVYQAACALLPDLEILPSSGRTIEGEGVNLSSGQKRVSLAR 780
Qy      781 AVYSNADITYLPDDPLSAVDHVGHIPEENVIGPRGMLKNKTRILVTHSMSTLPQVDYIIV 840
Db      781 AVYSNADITYLPDDPLSAVDHVGHIPEENVIGPRGMLKNKTRILVTHSMSTLPQVDYIIV 840
Qy      841 MSGGKISEMGSYQELRLARDGAFAFLRTYASTEQDAEENGVTGVSQPKGAQOMENGM 900
Db      841 MSGGKISEMGSYQELRLARDGAFAFLRTYASTEQDAEENGVTGVSQPKGAQOMENGM 900
Qy      901 LVYDSAGKOLQROLSSSSYSISGDISRHNSTAEIQAKBAKKEEFWKLMADKATQGVYKL 960
Db      901 LVYDSAGKOLQROLSSSSYSISGDISRHNSTAEIQAKBAKKEEFWKLMADKATQGVYKL 960

```

[illegible]

DR	EMBL; AF022838; AAB83980.1; JOINED; Genomic DNA.
DR	EMBL; AF022848; AAB83980.1; JOINED; Genomic DNA.
DR	EMBL; AF022847; AAB83980.1; JOINED; Genomic DNA.
DR	EMBL; AF022846; AAB83980.1; JOINED; Genomic DNA.
DR	EMBL; AF022845; AAB83980.1; JOINED; Genomic DNA.
DR	EMBL; AF022844; AAB83980.1; JOINED; Genomic DNA.
DR	EMBL; AF022843; AAB83980.1; JOINED; Genomic DNA.
DR	EMBL; AF022842; AAB83980.1; JOINED; Genomic DNA.
DR	EMBL; AF022841; AAB83980.1; JOINED; Genomic DNA.
DR	EMBL; AF022839; AAB83980.1; JOINED; Genomic DNA.
DR	EMBL; AF022852; AAB83980.1; JOINED; Genomic DNA.
DR	EMBL; AF022851; AAB83980.1; JOINED; Genomic DNA.
DR	EMBL; AF022850; AAB83980.1; JOINED; Genomic DNA.
DR	EMBL; AF022849; AAB83980.1; JOINED; Genomic DNA.
DR	EMBL; AF022837; AAB83980.1; JOINED; Genomic DNA.
DR	EMBL; AF022835; AAB83980.1; JOINED; Genomic DNA.
DR	EMBL; AF022833; AAB83980.1; JOINED; Genomic DNA.
DR	EMBL; AF022831; AAB83980.1; JOINED; Genomic DNA.
DR	EMBL; AF022829; AAB83980.1; JOINED; Genomic DNA.
DR	EMBL; AF022827; AAB83980.1; JOINED; Genomic DNA.
DR	HSSP; P08716; ILMT0.
DR	Ensembl; ENSG00000103222; Homo sapiens.
DR	GO; GO:0016021; C:integral to membrane; IEA.
DR	GO; GO:0005524; F:ATP binding; IEA.
DR	GO; GO:0016887; F:ATPase activity; IEA.
DR	GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . ; IEA.
DR	GO; GO:0000166; F:nucleotide binding; IEA.
DR	GO; GO:0006810; P:transport; IEA.
DR	InterPro; IPRO03593; AAA ATPase.
DR	InterPro; IPRO11527; ABC_membrane_1.
DR	InterPro; IPRO01140; ABC_TM_transp.
DR	InterPro; IPRO003439; ABC_transp_lfke.
DR	InterPro; IPRO05292; MRP_assoc.
DR	InterPro; IPRO00719; Prot_kinase.
DR	Pfam; PF00664; ABC_membrane_2.
DR	Pfam; PF00005; ABC_tran; 1.
DR	ProDom; PD000006; ABC_transporter; 1.
DR	SMART; SM00382; AAA; 2.
DR	TIGFAMS; TIGR00957; MRP_assoc_pro; 1.
DR	PROSITE; PS02929; ABC_TMlf_2; UNKNOWN_1.
DR	PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN_1.
DR	PROSITE; PS0893; ABC_TRANSPORTER_2; 2.
DR	PROSITE; PS0107; PROTEIN_KINASE_ATP; UNKNOWN_1.
KW	ATP-binding; Nucleotide-binding.
FT	NON TER
SQ	SEQUENCE 1459 AA; 163831 MW; 6A016A5AE7AA95C CRC64;
<hr/>	
Query Match	75.2%; Score 7444; DB 2; Length 1459;
Best Local Similarity	96.3%; Pred. No. 0;
Matches 1459; Conservative	0; Mismatches 0; Indels 56; Gaps 1;
<hr/>	
Gy	17 DMNVTMTNSNDPFKFCQNTVTVWPVCPLYMACFPFYLYSRHRDGIOWTPINKTKTA 76
Dd	1 DMMVTMTNSNDPFKFCQNTVTVWPVCPLYMACFPFYLYSRHRDGIOWTPINKTKTA 60
Gy	77 LGFLIMVICMADLIFYSEWESRGIFLAPVPLVSEPTLLGITTLATPFLQLERRKGVOSSG 136
Dd	61 LGLFLIMVICMADLIFYSEWESRGIFLAPVPLVSEPTLLGITTLATPFLQLERRKGVOSSG 120
Gy	137 IMLTFFMVALYCALAIIRSKITMLAKEDAQVDLFREDITFYVFSLILIIQLVLSCFSDRSP 196
Dd	121 IMLTFFMVALYCALAIIRSKITMLAKEDAQVDLFREDITFYVFSLILIIQLVLSCFSDRSP 180
Gy	197 LFSEITHDPNCPRESSASPLSRIITFMWITGILVNGVQPIEGSDMLSNKEDTSBOVVAV 256
Dd	181 LFSEITHDPNCPRESSASPLSRIITFMWITGILVNGVQPIEGSDMLSNKEDTSBOVVAV 240
Gy	257 LVKNMKKECATRTQPKVKVYVSXKDPAQPKSSKYDNANEEVEALIVKSPOKENMPSLFKV 316
Dd	241 LVKNMKKECATRTQPKVKVYVSXKDPAQPKSSKYDNANEEVEALIVKSPOKENMPSLFKV 300
Gy	317 LYKTGGVYFLMSPPFKAIHDLNMFSGQIILKLIKFYNDITKADPWOGYFTYTVLLFTVACL 376
Dd	317 LYKTGGVYFLMSPPFKAIHDLNMFSGQIILKLIKFYNDITKADPWOGYFTYTVLLFTVACL 376

Db 301 LYKTEGFEYFLMSFEFKAIHDLMMFSGPQIILKLIKFVNDTKAPDMOGYFYTVLLFVTACL 360
Qy 377 QTLVHOYFPHICFSGMGRIKTAVIGAVYRKALVTITNSARKSSTGEIYNLMSVDAQSPMD 436
Db 361 QTLVHOYFPHICFSGMGRIKTAVIGAVYRKALVTITNSARKSSTGEIYNLMSVDAQSPMD 420
Qy 437 LATVNMWSPAPLOVILALYLLMLNLTGSPVLAVGVAVMVLVAVNAVMAMKTKTYQVAHMK 496
Db 421 LATVNMWSPAPLOVILALYLLMLNLTGSPVLAVGVAVMVLVAVNAVMAMKTKTYQVAHMK 480
Qy 497 SKDNRIKLMNEIILNGIKYKLYAMELAFKDYLAIRQELKVKLKSAYLSVGTFTWCT 556
Db 481 SKDNRIKLMNEIILNGIKYKLYAMELAFKDYLAIRQELKVKLKSAYLSVGTFTWCT 540
Qy 557 PFLVALCTFAVYVITIDENNIIIDAOTAFVSLFNLIRPFLNITLPMVSIISYQASVSLKRL 616
Db 541 PFLVALCTFAVYVITIDENNIIIDAOTAFVSLFNLIRPFLNITLPMVSIISYQASVSLKRL 600
Qy 617 RIFLSHEELEBPSIERRPVKGGGTSITVNNATFTMARSDPPTLNGITFSIPGALVAV 676
Db 601 RIFLSHEELEBPSIERRPVKGGGTSITVNNATFTMARSDPPTLNGITFSIPGALVAV 660
Qy 677 VGQVCGCKSSLLSLALAMDVEGHVAKGSVAVYPOQAMTQNDISRENTILFGQLEBPY 736
Db 661 VGQVCGCKSSLLSLALAMDVEGHVAKGSVAVYPOQAMTQNDISRENTILFGQLEBPY 720
Qy 737 YRSYIOACALLPDEIILPSGDRTEIGEKGVMLSGGQKQVSLARAVYSNADIYLFDDPLS 796
Db 721 YRSYIOACALLPDEIILPSGDRTEIGE----- 747
Qy 797 AVDAHVGKHIFENYIGPKGMLKNKTRILVTHSMSYLPOVDYIIVMSGKISEMGSYQELL 856
Db 748 -----KTRILVTHSMSYLPQVDYIIVMSGKISEMGSYQELL 784
Qy 857 ARDQAPAFELRTYASTEBODAEENGVTGSGPGKEAQMENGMLVTBSAKQORQSS 916
Db 785 ARDQAPAFELRTYASTEBODAEENGVTGSGPGKEAQMENGMLVTBSAKQORQSS 844
Qy 917 SSSYSGDISRRHNSTAELOKAEAKKEFTWKLEADKAQTGVKLSVYDYMKAIGLFISSF 976
Db 845 SSSYSGDISRRHNSTAELOKAEAKKEFTWKLEADKAQTGVKLSVYDYMKAIGLFISSF 904
Qy 977 LSIPLFMCNHNVSALASNWYLSLMTDDPIVNGTOEHKTVRLSVYGAIGISOGIAVFGYMA 1036
Db 905 LSIPLFMCNHNVSALASNWYLSLMTDDPIVNGTOEHKTVRLSVYGAIGISOGIAVFGYMA 964
Qy 1037 VSIQGIILASRCLHVDLHSIIRSPMSFFERTPSGNLVNRFSEKELDTVDSMTPEVTKMFG 1096
Db 965 VSIQGIILASRCLHVDLHSIIRSPMSFFERTPSGNLVNRFSEKELDTVDSMTPEVTKMFG 1024
Qy 1097 SLFNVIGACIYILATPIAAIIPPLGLIYFFVORFFVYASSRQRLKLSVSRSPIYSHFN 1156
Db 1025 SLFNVIGACIYILATPIAAIIPPLGLIYFFVORFFVYASSRQRLKLSVSRSPIYSHFN 1084
Qy 1157 ETLGVSVYIRAFEEQERFIHOSDLKVDENOKAYYSIIVANRWLAARLECVGNCILFPAAL 1216
Db 1085 ETLGVSVYIRAFEEQERFIHOSDLKVDENOKAYYSIIVANRWLAARLECVGNCILFPAAL 1144
Qy 1217 FAVISRHSLSAGLVGLSVYSIQVTTYLANVLVRSSSEMETNIVAVERLKEYSETEKEAPW 1276
Db 1145 FAVISRHSLSAGLVGLSVYSIQVTTYLANVLVRSSSEMETNIVAVERLKEYSETEKEAPW 1204
Qy 1277 QIOETAPSSMPQVGRVEFRNYCLAYRBDLPVLHINVTIINGKRGVIGRGTKGKSSL 1336
Db 1205 QIOETAPSSMPQVGRVEFRNYCLAYRBDLPVLHINVTIINGKRGVIGRGTKGKSSL 1264
Qy 1337 TLGFRINESABGEIILIDGINIAKIGLHDLRKITIIIPQDPVLPFGSGSLRMLNDPPSOYSD 1396
Db 1265 TLGFRINESABGEIILIDGINIAKIGLHDLRKITIIIPQDPVLPFGSGSLRMLNDPPSOYSD 1324
Qy 1397 BEVWTSLELAHKDFVSALPDKLDHECABGGENISVGOQOLVCLARALLRKTKIILVDEBA 1456
Db 1325 BEVWTSLELAHKDFVSALPDKLDHECABGGENISVGOQOLVCLARALLRKTKIILVDEBA 1384

Qy 1457 TAAVDETDLLIGSTIRTOPEDCVTLTAHRLNTIMDYTRVYILDKGEIOYGAPSDLLQ 1516
Db 1385 TAAVDETDLLIGSTIRTOPEDCVTLTAHRLNTIMDYTRVYILDKGEIOYGAPSDLLQ 1444
Qy 1517 ORGLFYSMAQDAGLV 1531
Db 1445 ORGLFYSMAQDAGLV 1459
RESULT 6
QY09040_HUMAN
ID Q9040_HUMAN PRELIMINARY; PRT; 1456 AA.
AC Q9040;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 25, Last annotation update)
DE Multidrug resistance protein (Fragment).
GN Name=MRP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;
OC Homo.
NCBI_TaxID=9606;
RX NCBI
RP NCBI
RP NCBI
RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;
RA Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;
RT "Analysis of the intron-exon organization of the human multidrug-
resistance protein gene (MRP) and alternative splicing of its mRNA.";
RL Genomics 45:368-378(1997).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF022853; AAB83981.1; -; Genomic DNA.
DR EMBL; AF022824; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022825; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022826; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022827; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022829; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022831; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022833; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022835; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022837; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022847; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022846; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022845; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022844; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022843; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022842; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022841; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022840; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022838; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022852; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022851; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022850; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022849; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022848; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022836; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022834; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022832; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022830; AAB83981.1; JOINED; Genomic DNA.
DR EMBL; AF022828; AAB83981.1; JOINED; Genomic DNA.
DR HSP; P08716; IMTO.
DR Ensembl; ENSG00000103222; Homo sapiens.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transep_like.

DR InterPro: IPR005292; MRP assoc.
 DR InterPro: IPR000719; Prot. kinase.
 DR Pfam: PF00664; ABC membrane; 2.
 DR Pfam: PF00005; ABC_tran; 2.
 DR ProDom: PD000006; ABC_transporter; 2.
 DR SMART: SM00382; AAA; 2.
 DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
 DR PROSITE: PS00929; ABC_TM1F; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
 DR PROSITE: PS00893; ABC_TRANSPORTER_2; 2.
 DR PROSITE: PS01071; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR ATP-binding; Nucleotide-binding.
 FT NON TER 1
 SQ SEQUENCE 1456 AA; 163233 MW; 8DB8AAB2BC481F2 CRC64;

Query Match 74.9%; Score 7419.5; DB 2; Length 1456;
 Best Local Similarity 96.1%; Pred. No. 0;
 Matches 1456; Conservative 0; Mismatches 0; Indels 59; Gaps 1;

QY 17 DAWVTMTSNPDFTKCFONTVLVWPCFYLACFPYFLYLRDRGTYQMTPLNKTXTA 76
 DB 1 DAWVTMTSNPDFTKCFONTVLVWPCFYLACFPYFLYLRDRGTYQMTPLNKTXTA 60

QY 77 LGFLMTIVCVDLFFYSFMRSGIFLAPVPLVSPFLIGITTLATFLIQERRKGVSSG 136
 DB 61 LGFLMTIVCVDLFFYSFMRSGIFLAPVPLVSPFLIGITTLATFLIQERRKGVSSG 120

QY 137 IMTLFMTVALYCALAIRSKMTMLKEDAOVDLFRDITFYVFSLLLIQLVLCSPDRSP 196
 DB 121 IMTLFMTVALYCALAIRSKMTMLKEDAOVDLFRDITFYVFSLLLIQLVLCSPDRSP 180

QY 197 LFSETIHDPNCPSSASPLSRITFMWITGLIVRGYOPLEGSDLMSEKEDTSEOVVPV 256
 DB 181 LFSETIHDPNCPSSASPLSRITFMWITGLIVRGYOPLEGSDLMSEKEDTSEOVVPV 240

QY 257 LVKNMKCEAKTRKQPKVYVSSKDPAPCKESSKYDANEVEALIVKSPCKEMPSLFKV 316
 DB 241 LVKNMKCEAKTRKQPKVYVSSKDPAPCKESSKYDANEVEALIVKSPCKEMPSLFKV 300

QY 317 LYKTGGYFLMSFFPKAIHDLMSFGQILKLIKPVNDTKAPDQGYFTYVLLFTVACL 376
 DB 301 LYKTGGYFLMSFFPKAIHDLMSFGQILKLIKPVNDTKAPDQGYFTYVLLFTVACL 360

QY 377 QTLVLHGYPHICFVSGRIKTAIVGAYRRKALVITNSARKSTGELIVNLSVDAQSPMD 436
 DB 361 QTLVLHGYPHICFVSGRIKTAIVGAYRRKALVITNSARKSTGELIVNLSVDAQSPMD 420

QY 437 LATYINMISAPLQVILALYLLMLNLGPSVLAVAVWVAVVAVVAVVAVVAVVAVVAVV 496
 DB 421 LATYINMISAPLQVILALYLLMLNLGPSVLAVAVVAVVAVVAVVAVVAVVAVVAVV 480

QY 497 SKDNRIKLMEIILNGIVLKIYAMELAFKDKVLAIRQBELKVLKKSAYLSAVGTFTWCT 556
 DB 481 SKDNRIKLMEIILNGIVLKIYAMELAFKDKVLAIRQBELKVLKKSAYLSAVGTFTWCT 540

QY 557 PFLVALCTEAVVYVITDENNILLDAQTAFLALPFLILPFLILPFLILPFLILPFLILP 616
 DB 541 PFLVALCTEAVVYVITDENNILLDAQTAFLALPFLILPFLILPFLILPFLILPFLILP 600

QY 617 RIFLSHELSPDISIERRPVKGDTNSITVYNAFTVARSQDPTLNGITISIPGALVAV 676
 DB 601 RIFLSHELSPDISIERRPVKGDTNSITVYNAFTVARSQDPTLNGITISIPGALVAV 660

QY 677 VGVGVCKSSLSALLAEMDKVEGHVAIKGSVAVYVPOAMTQNDLRENIILFGQCLEEPY 736
 DB 661 VGVGVCKSSLSALLAEMDKVEGHVAIKGSVAVYVPOAMTQNDLRENIILFGQCLEEPY 688

QY 737 YRSVYQACALLPDLLEIPSGDRTEIGKGVNLSSGCKORVSIARAVYSNADIYLFDDPLS 796
 DB 689 YRSVYQACALLPDLLEIPSGDRTEIGKGVNLSSGCKORVSIARAVYSNADIYLFDDPLS 721

QY 797 AVDAHVGHKIFENVIGPKMLKNKTRILVTHSMGYLPQVAVIIVMSGKISSEMGSYOELL 856
 DB 797 AVDAHVGHKIFENVIGPKMLKNKTRILVTHSMGYLPQVAVIIVMSGKISSEMGSYOELL 856

DB 722 AVDAHVGHKIFENVIGPKMLKNKTRILVTHSMGYLPQVAVIIVMSGKISSEMGSYOELL 781

QY 857 ARDGAFAEPLRTVASTQEQDAEENGYTVSGPCKEAKOMENGLVYDSAGKQIQROLSS 916
 DB 782 ARDGAFAEPLRTVASTQEQDAEENGYTVSGPCKEAKOMENGLVYDSAGKQIQROLSS 841

QY 917 SSSYSGDISRHNNSTAELOKAEEKETWKLMEADKQOTGOVKLSVYWDYKAIGLFISF 976
 DB 842 SSSYSGDISRHNNSTAELOKAEEKETWKLMEADKQOTGOVKLSVYWDYKAIGLFISF 901

QY 977 LSTFLFMCNVSALASNYMLSLMTDDEIYNGTOEHTKVRLSVYGALGISQIAVFGYSMA 1036
 DB 902 LSTFLFMCNVSALASNYMLSLMTDDEIYNGTOEHTKVRLSVYGALGISQIAVFGYSMA 961

QY 1037 VSTIGILASRCINHDLSILSRPMSFEETPSGNLVNRSKELDTYDSMPEYIKNPMG 1096
 DB 962 VSTIGILASRCINHDLSILSRPMSFEETPSGNLVNRSKELDTYDSMPEYIKNPMG 1021

QY 1097 SLFNVIGACIVILLATPIAIIIPPLGLIYFVQRFVVASRQLKLESYSRSPVYSHFN 1156
 DB 1022 SLFNVIGACIVILLATPIAIIIPPLGLIYFVQRFVVASRQLKLESYSRSPVYSHFN 1081

QY 1157 ETLIGSVVIRAFEBQERFIHQSDLKVDENQKAYYPSIVANRWLAVRLECVGNCTVLPAL 1216
 DB 1082 ETLIGSVVIRAFEBQERFIHQSDLKVDENQKAYYPSIVANRWLAVRLECVGNCTVLPAL 1141

QY 1217 FAYISRHSLSAGLVGSLVSYSLQVTTYLMNLVNMSSMEFNIVAVERLKSYSEFEKAPW 1276
 DB 1142 FAYISRHSLSAGLVGSLVSYSLQVTTYLMNLVNMSSMEFNIVAVERLKSYSEFEKAPW 1201

QY 1277 QIOETAPSSMPQVGRFEPFNVCYLRVEDLDPVLRHINVTINGEKXGIVRTAGASSL 1336
 DB 1202 QIOETAPSSMPQVGRFEPFNVCYLRVEDLDPVLRHINVTINGEKXGIVRTAGASSL 1261

QY 1337 TLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLFSSGLRMLNDFPSQYSD 1396
 DB 1262 TLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTIIPQDPVLFSSGLRMLNDFPSQYSD 1321

QY 1397 EEWYTSLELAHLKDPFYSALPDKLDHBCAEGENLSVQORLVCLARALLRKTKILVIDEA 1456
 DB 1322 EEWYTSLELAHLKDPFYSALPDKLDHBCAEGENLSVQORLVCLARALLRKTKILVIDEA 1381

QY 1457 TAAVDLETDDLIOSTIRTOPEDCVLTIAHRLMTIMRYEVIYLDKKEIOEGAPSLQ 1516
 DB 1382 TAAVDLETDDLIOSTIRTOPEDCVLTIAHRLMTIMRYEVIYLDKKEIOEGAPSLQ 1441

QY 1517 ORGLFYSMADAGLV 1531
 DB 1442 ORGLFYSMADAGLV 1456

RESULT 7
 Q6UR05 CANFA PRELIMINARY; PRT; 1531 AA.
 ID O6UR05- CANFA PRELIMINARY;
 AC O6UR05-
 DT 05-JUL-2004 (Tremblere1. 27, Created)
 DT 05-JUL-2004 (Tremblere1. 27, last sequence update)
 DT 05-JUL-2004 (Tremblere1. 27, last annotation update)
 DE Multidrug resistance-associated protein 1.
 GN Name=MRP1;
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 OC Canis.
 OC NCBI_TaxID=9615;
 RN [1]
 RP NCBIOTIDE SEQUENCE.
 RX MEDLINE=22404678; PubMed=12516967;
 RA Ma L., Pratt S.B., Cao J., Dantzig A.H., Moore R.E., Slapak C.A.;
 RT "Identification and characterization of the canine multidrug
 resistance-associated protein.";
 RL Mol. Cancer Ther. 1:1335-1342(2002).
 RN [2]

RP NUCLEOTIDE SEQUENCE.
RA Ma L., Pratt S.E., Cao J., Danczig A.H., Moore R.E., Slapak C.A.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL: AY63728; AAC3348.1; mRNA.
DR EMBL: ENSCAFG0000018208; Canis familiaris.
DR GO: GO:0016021; C:Integral to membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0042666; F:ATPase activity; IEA.
DR GO: GO:0000166; F:nucleotide binding; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR003593; AAA ATPase.
DR InterPro: IPR011527; ABC membrane 1.
DR InterPro: IPR01140; ABC TM transp.
DR InterPro: IPR003439; ABC transp_like.
DR InterPro: IPR005292; MRP_aasoc.
DR InterPro: IPR00719; Prot_kinase.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD000006; ABC_transporter; 2.
DR SMART: SMO0382; AAA; 2.
DR TIGRfams: TIGR00957; MRP_aasoc_pro; 1.
DR PROSITE: PS0929; ABC_TM1; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE: PS0893; ABC_TRANSPORTER_2; 2.
DR PROSITE: PS0107; PROTEIN KINASE_ATP; UNKNOWN 1.
DR ATP-binding: Nucleotide-binding; Repeat: Transport.
KW SEQUENCE 1531 AA; 171792 MW; 746361A71C6158BD CRC64;

Query Match 73.4%; Score 7272; DB 2; Length 1531;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 1402; Conservative 70; Mismatches 59; Indels 0; Gaps 0;

QY 1 MALSFGSADSDSLMDNMNTNMTNPNDFKCFQNTVYVWYPCFYLMACFPFFLYLSRH 60
DB 1 MALSFGSADSDSLPFWMDVSMNTSNPDTKCFQNTVYVWVPCCLMCFPFYLYLSRH 60
QY 61 DRGYIOMTPLNKRTALGFLMIWCMADLFYSFWERSRGIPLAPYFLVSPITLIGITTLA 120
DB 61 DRGYIOMTPLNKRTALGFLMIWCMADLFYSFWERSRGIPLAPYFLVSPITLIGITTLA 120
QY 121 TFLIQLEERRKGVSSGIMLTFWLVALVCALILRSKIMTALKEDAQVDLPFDITFYVYS 180
DB 121 TFLIQLEERRKGVSSGIMLTFWLVALVCALILRSKIMTALKEDAQVDLPFDITFYVYS 180
QY 181 LLLIQVLVSCFSDSPLEFSETIHDNPPCBSSASLSITTFWITGLVGRYGRPLESSD 240
DB 181 LVLQLVLVSCFSDSPLEFSETIHDNPPCBSSASLSITTFWITGLVGRYGRPLESSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNMKECAKTRKQPVKVYVSSKDPAPCKESSKYDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNMKECAKTRKQPVKVYVSSKDPAPCKESSKYDANEVEAL 300
QY 301 IVKSPQKEMNSLFFKVLVYKTFGPFYFLMSFFKAIHDLMMFSGPOLIKLIFVNDTKAP 360
DB 301 IVKTPQKREBSLFFKVLVYKTFGPFYFLMSFFKAIHDLMMFSGPOLIKLIFVNDTKAP 360
QY 361 WQGFYTYLLFVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 420
DB 361 WQGFYTYLLFVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 420
QY 421 GEIYVLSVDAORFMDLATYINMISAPLOVTLAYLLMLNGSPVLGAVVMVLMVNVN 480
DB 421 GEIYVLSVDAORFMDLATYINMISAPLOVTLAYLLMLNGSPVLGAVVMVLMVNVN 480
QY 481 AVNMAKTKTYOVANMKSXONRIKLMNEILNGIKVLKLYAMBLAFKDYALROBELKYLK 540
DB 481 AVNMAKTKTYOVANMKSXONRIKLMNEILNGIKVLKLYAMBLAFKDYALROBELKYLK 540
QY 541 KSAIYLSAVGTWCTPPLVALCTPAVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 600
DB 541 KSAIYLSAVGTWCTPPLVALCTPAVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYV 600

QY 601 MVISSIVQASVSLKRLRIFLSHELEBDSIERRPVKDGGGINSITVNNATFTWASDPPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHELEBDSIERRPVKDGGGINSITVNNATFTWASDPPT 660
QY 661 LMGITTSIPGALVAVVGVQCGKSSLSLLALMDKYBEHVAKGSVAVYVPOQAWIQND 720
DB 661 LSGITTSIPGSLVAVVGVQCGKSSLSLLALMDKYBEHVAKGSVAVYVPOQAWIQND 720
QY 721 SLRENIIFGQLEPPRYRSVYIACALLPDEIIPSGDRTGEGKVMSSGQKORVSLAR 780
DB 721 SLRENIIFGQLEPPRYRSVYIACALLPDEIIPSGDRTGEGKVMSSGQKORVSLAR 780
QY 781 AVYSNADIIYFDDPESAVDAHVGNHIFENYIGPKMKLNKTRILVYTHSMSTLPQVDYIIV 840
DB 781 AVYCDSDIYFDDPESAVDAHVGNHIFENYIGPKMKLNKTRILVYTHSMSTLPQVDYIIV 840
QY 841 MSGKISEMSGYQELRLARDGAPAFLLTYASTBOEDAEENGVTGVSQPKAKOMENGM 900
DB 841 MTGKISEMSGYQELRLARDGAPAFLLTYASTBOEDAEENGVTGVSQPKAKOMENGM 900
QY 901 LVYDSAGKOLQORLSSSSYSQDTSRHHNSTAELQKAEAKKEBTWKLMEADKATGVYKL 960
DB 901 LVYDSAGKOLQORLSSSSYSQDTSRHHNSTAELQKAEAKKEBTWKLMEADKATGVYKL 960
QY 961 SVYDWYKAKLGLFISFLSIFLFCMCHVASALASNYMSLWTDPIVNGTQEHTRVLSVYG 1020
DB 961 SVYDWYKAKLGLFISFLSIFLFCMCHVASALASNYMSLWTDPIVNGTQEHTRVLSVYG 1020
QY 1021 ALGISOGIAPVGSMAVSIIGIILASRCLHYDLHSILRSMPFPERPSGULVNRFSKEL 1080
DB 1021 ALGISOGIAPVGSMAVSIIGIILASRCLHYDLHSILRSMPFPERPSGULVNRFSKEL 1080
QY 1081 PTVDSMIPVYKMGMSLFNVYIACIYLLATPLAIIIPPLGIYFFVQRFYVASSROL 1140
DB 1081 PTVDSMIPVYKMGMSLFNVYIACIYLLATPLAIIIPPLGIYFFVQRFYVASSROL 1140
QY 1141 KRLSEVSRSPPVSHFNSTLLGVSVYIRAFEEQEREIHOSSDKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLSEVSRSPPVSHFNSTLLGVSVYIRAFEEQEREIHOSSDKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLCEVNCIYLPALFAVIRSHSLAGVLVSYSLOVTTYVNMIVRSMSEMETNIVA 1260
DB 1201 VRLCEVNCIYLPALFAVIRSHSLAGVLVSYSLOVTTYVNMIVRSMSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIQENAPSPWPQVGRVAFRDIYGLRYKRENDLVLAHINITINGG 1320
DB 1261 VERLKEYSETEKEAPWQIQENAPSPWPQVGRVAFRDIYGLRYKRENDLVLAHINITINGG 1320
QY 1321 EKVGIVRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIVRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGIIRPMNIDPPSOYSDDEWVTSLELAHLKDFVSALPKLDHECAEGENSVGQROLVCL 1440
DB 1381 SGIIRPMNIDPPSOYSDDEWVTSLELAHLKDFVSALPKLDHECAEGENSVGQROLVCL 1440
QY 1441 ARALLRRTKTIIVDEAPRAVDLETFDILLQSTIRQFPDCTVLTIAHRLNTIMDTRYIVL 1500
DB 1441 ARALLRRTKTIIVDEAPRAVDLETFDILLQSTIRQFPDCTVLTIAHRLNTIMDTRYIVL 1500
QY 1501 DKGEIOEGAPSDILOQRLGFYSMAKDAGLV 1531
DB 1501 DKGEIRREGQPSDILQQRGLFYSMAKDAGLV 1531

RESULT 8
Q59G19 HUMAN
ID Q59G19 HUMAN PRELIMINARY; PRT; 1439 AA.
AC Q59G19
DT 10-MAY-2005 (TRENBLrel. 30, Created)
DT 10-MAY-2005 (TRENBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TRENBLrel. 30, Last annotation update)

DE ATP-binding cassette, sub-family C, member 1 isoform 1 variant
(Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain.
RA Tsuchi Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.,
RT "None Title."
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL: AB009120; BAD92357.1; -; mRNA.
DR Ensembl: ENSG00000103222; Homo sapiens.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:005524; F:ATP binding; IEA.
DR GO: GO:0016887; F:ATPase activity; IEA.
DR GO: GO:000166; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR011527; ABC_membrane_1.
DR InterPro: IPR001140; ABC_TM_transp.
DR InterPro: IPR003439; ABC_transp_like.
DR InterPro: IPR005292; MRP_assoc.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR SMART: SM00382; AAA; 2.
DR TIGRfam: TIGR00957; MRP_assoc_pro; 1.
DR PROSITE: PS00929; ABC_TMIF; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE: PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR ATP-binding; Nucleotide-binding.
KW NON_TER
FT
SQ SEQUENCE 1439 AA; 160439 MW; 090D881ADBFB3077 CRC64;

Query Match 73.1%; Score 7242; DB 2; Length 1439;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1428; Conservative 0; Mismatches 1; Indels 10; Gaps 1;

QY 103 APVFLVSPFLIGITTLATFLIQLERRKGVQSSGIMLTFMLVALVCALALRSKIMTALK 162
DB 1 APVFLVSPFLIGITTLATFLIQLERRKGVQSSGIMLTFMLVALVCALALRSKIMTALK 60

QY 163 EDNOVLFRLDITFYVYFSLILLQVLVSCFSDRSPLPSETIHDNPPCESASFLSRITFW 222
DB 61 EDNOVLFRLDITFYVYFSLILLQVLVSCFSDRSPLPSETIHDNPPCESASFLSRITFW 120

QY 223 WINGLIVRGYRQPLEGSDMLSNKEPTSEOVVPLVYKMKKECAKTRKQVYKVVYSSKDP 282
DB 223 WINGLIVRGYRQPLEGSDMLSNKEPTSEOVVPLVYKMKKECAKTRKQVYKVVYSSKDP 160

QY 121 WITGLIVRGYRQPLEGSDMLSNKEPTSEOVVPLVYKMKKECAKTRKQVYKVVYSSKDP 160
DB 121 WITGLIVRGYRQPLEGSDMLSNKEPTSEOVVPLVYKMKKECAKTRKQVYKVVYSSKDP 160

QY 283 AQRKSSKVDANEVEALIVKSPQKEMNPSLFVLVYKTFGPFYLMSEFFFAIHDLMVFSG 342
DB 181 AQRKSSKVDANEVEALIVKSPQKEMNPSLFVLVYKTFGPFYLMSEFFFAIHDLMVFSG 240

QY 343 POLIKLILKRVNNTKAPDMNGFYFYVLLPTACIQTLVLHQRHICVSGMRKTAIVIGA 402
DB 241 POLIKLILKRVNNTKAPDMNGFYFYVLLPTACIQTLVLHQRHICVSGMRKTAIVIGA 300

QY 403 VYKALVITNSARKSSIVGEIVNIMSVDAQRFMDLATYINMISAPLOVILLALVILMLN 462
DB 301 VYKALVITNSARKSSIVGEIVNIMSVDAQRFMDLATYINMISAPLOVILLALVILMLN 360

QY 463 GRSVLAGVAVMLVMPNVAVMAMKTKTYQVAHKMSKDNRIKLMNELLNGIKVILKYAMEL 522
DB 361 GRSVLAGVAVMLVMPNVAVMAMKTKTYQVAHKMSKDNRIKLMNELLNGIKVILKYAMEL 420

QY 523 AFKDKVLAHQEELKVLKKSAYLSAVGTFPTWCTPFLVALCTFAVYVITDENNLDAQTA 582
DB 421 AFKDKVLAHQEELKVLKKSAYLSAVGTFPTWCTPFLVALCTFAVYVITDENNLDAQTA 480

QY 583 FVSLALFNILRPNLILPWLIVSIIVASVSLKRLIFLSHELEPDSIERRPVQGGGTN 642
DB 481 FVSLALFNILRPNLILPWLIVSIIVASVSLKRLIFLSHELEPDSIERRPVQGGGTN 540

QY 643 SITVRNATFTWASDDPTLNGITFSPLEGALVAVVQVGGCKSSLSALLAENDKRGHV 702
DB 541 SITVRNATFTWASDDPTLNGITFSPLEGALVAVVQVGGCKSSLSALLAENDKRGHV 600

QY 703 AIKGSVAVYVQOAMQNDLSRENILFGCCLEBPYYSVIOACALLPLLEILPSGDRFEG 762
DB 601 AIKGSVAVYVQOAMQNDLSRENILFGCCLEBPYYSVIOACALLPLLEILPSGDRFEG 660

QY 763 EKGVLNCGGQKQVSLARAVYNSADITLPPDPLSAVDHVGKHI FENVIGPKGMKNKTR 822
DB 661 EKGVLNCGGQKQVSLARAVYNSADITLPPDPLSAVDHVGKHI FENVIGPKGMKNKTR 720

QY 823 ILVTHSMSTYLPQVDVITVNSGGKISMGYSQELIARDGAFELRTYASTEOBDAEEN 881
DB 721 ILVTHSMSTYLPQVDVITVNSGGKISMGYSQELIARDGAFELRTYASTEOBDAEEN 780

QY 882 -----GVTVSGPKGKAKOMENGMLVTDSAGKOLQRLSSSSSGSDISRHNSTA 932
DB 781 STWDEBEAGVTVSGGKAKOMENGMLVTDSAGKOLQRLSSSSSGSDISRHNSTA 840

QY 933 ELQKAEAKKEETKMLEADKAQTGYVLASYMDYMKALGLIFISLIFLPMCHVSLAS 992
DB 841 ELQKAEAKKEETKMLEADKAQTGYVLASYMDYMKALGLIFISLIFLPMCHVSLAS 900

QY 993 NYWLSLMTDDPIVNGTQEHKTVLSYVGALGISQGLAVFGYSMAVSGTILASRCLHVL 1052
DB 901 NYWLSLMTDDPIVNGTQEHKTVLSYVGALGISQGLAVFGYSMAVSGTILASRCLHVL 960

QY 1053 LHSILSPMSFFERTSSGNLVNRFSEKELDVDSMIEBVIKMFNGSLPNVGACIVILLAT 1112
DB 961 LHSILSPMSFFERTSSGNLVNRFSEKELDVDSMIEBVIKMFNGSLPNVGACIVILLAT 1020

QY 1113 PIAATIIIPGLIYFPQRFVYVASSRQLKRLSESVSRSPVYSHNETLIGVSVIRAFPEO 1172
DB 1021 PIAATIIIPGLIYFPQRFVYVASSRQLKRLSESVSRSPVYSHNETLIGVSVIRAFPEO 1080

QY 1173 RFIHQSDLVKDENOKAYYSIVANRWLAVERLCEVGNICVLPALPAVISHSLSAGLV 1232
DB 1081 RFIHQSDLVKDENOKAYYSIVANRWLAVERLCEVGNICVLPALPAVISHSLSAGLV 1140

QY 1233 SVSYSLQVTTYLNMVVRMSEMETNIVAVBRLEKYEETKEAWQIOETAPSPSSQVGR 1292
DB 1141 SVSYSLQVTTYLNMVVRMSEMETNIVAVBRLEKYEETKEAWQIOETAPSPSSQVGR 1200

QY 1293 VEFPRNYCLARYEDLDVLRHINVTINGEKVGIYGRTGAKSSLTGLFPIINSABEII 1352
DB 1201 VEFPRNYCLARYEDLDVLRHINVTINGEKVGIYGRTGAKSSLTGLFPIINSABEII 1260

QY 1353 IDGINIAKI GLHDLRFKITIIPQDPVLFSGSLRMNLDPSQYSDEEVMWLSLELAHKDPV 1412
DB 1261 IDGINIAKI GLHDLRFKITIIPQDPVLFSGSLRMNLDPSQYSDEEVMWLSLELAHKDPV 1320

QY 1413 SALPDLDBHECAEGGNSLVGQRLVCLARALLRKRLIVLDBATAVNDLETTDLIOSTI 1472
DB 1321 SALPDLDBHECAEGGNSLVGQRLVCLARALLRKRLIVLDBATAVNDLETTDLIOSTI 1380

QY 1473 RTQFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEGAFSDLLQOGLFFSMAKDAGLV 1531
DB 1381 RTQFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEGAFSDLLQOGLFFSMAKDAGLV 1439

RESULT 9
Q8HX05_BOVIN
ID Q8HX05_BOVIN PRELIMINARY; PRT; 1530 AA.
AC Q8HX05;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Multidrug resistance protein 1.
GN Name=MRP1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22062955; PubMed=12067707; DOI=10.1016/S0014-5793(02)02816-8;
RA Taguchi Y., Saeki K., Komano T.;
RT "Functional analysis of MRP1 cloned from bovine.";
RL FEBS Lett. 521:211-213(2002).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL: AB082124; BAC15550.1; -, mRNA.
DR HSP; P08716; IMT0.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0016887; F:ATPase activity; IEA.
DR GO: GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:0006810; F:transport; IEA.
DR InterPro: IPR035593; AAA_ATPase.
DR InterPro: IPR011527; ABC_membrane_1.
DR InterPro: IPR001140; ABC_TM_transp.
DR InterPro: IPR003439; ABC_transp_like.
DR InterPro: IPR005292; MRP_assoc.
DR Pfam: Pf00664; ABC_membrane; 2.
DR Pfam: Pf00005; ABC_tran; 2.
DR ProDom: PD000006; ABC_transporter; 2.
DR SMART: SM00382; AAA; 2.
DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
DR PROSITE: PS50929; ABC_TMIF; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE: PS50893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1530 AA; 171666 MW; AAE4F92ED7832703 CRC64;

Query Match 73.0%; Score 7230.5; DB 2; Length 1530;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1388; Conservative 83; Mismatches 59; Indels 1; Gaps 1;

QY 1 MALRGFSADGSDPLMDMNNVTWNTSNPFTKCFQNTVAVWPCFLIMACPEFFLYLSRH 60
DB 1 MALMDPFSVDSSDLFMENNVWNTSNPFTKCFQNTVAVWPCSTLWVCEFFFLYLSHH 60
QY 61 DRGYIQMTPLNKKTALGFLIMIVCWADLFYSFWEKRGIFLAPVFLVSPTLGITTLA 120
DB 61 DRGYIQMTPLNKKTALGFLIMIVCWADLFYSFWEKRGIFLAPVFLVSPTLGITTLA 120
QY 121 TFLIQLERRKGVSSGIMLTWLVAVCAALINSKIMTALKEDAVDLPDITFYVES 180
DB 121 TFLIQLERRKGVSSGIMLTWLVAVCAALINSKIMTALKEDAVDLPDITFYVES 180
QY 121 TFLIQLERRKGVSSGIMLTWLVAVCAALINSKIMTALKEDAVDLPDITFYVES 180
DB 121 TFLIQLERRKGVSSGIMLTWLVAVCAALINSKIMTALKEDAVDLPDITFYVES 180
QY 181 LLLIQVLVSCSDSPLEFSETHDNPCESSASFLSITFMWITGLVRYRQPLESSD 240
DB 181 LLLIQVLVSCSDSPLEFSETHDNPCESSASFLSITFMWITGLVRYRQPLESSD 240
QY 181 LLLIQVLVSCSDSPLEFSETHDNPCESSASFLSITFMWITGLVRYRQPLESSD 240
DB 181 LLLIQVLVSCSDSPLEFSETHDNPCESSASFLSITFMWITGLVRYRQPLESSD 240
QY 241 LMSLNKEDTSBOVAVLVKMKKECAKTRKQPVKVYVSSKDPAPKESKVDANEVEAL 300
DB 241 LMSLNKEDTSBOVAVLVKMKKECAKTRKQPVKVYVSSKDPAPKESKVDANEVEAL 300
QY 241 LMSLNKEDTSBOVAVLVKMKKECAKTRKQPVKVYVSSKDPAPKESKVDANEVEAL 300
DB 241 LMSLNKEDTSBOVAVLVKMKKECAKTRKQPVKVYVSSKDPAPKESKVDANEVEAL 300
QY 301 IVKSPQKEWNSLFRKVLVYKTFGPFVLMSPFFKAHIDLMFSGPOLIKLIRVNDTAPD 360
DB 301 IVKSPQKEWNSLFRKVLVYKTFGPFVLMSPFFKAHIDLMFSGPOLIKLIRVNDTAPD 360
QY 301 IVKSPQKEWNSLFRKVLVYKTFGPFVLMSPFFKAHIDLMFSGPOLIKLIRVNDTAPD 360
DB 301 IVKSPQKEWNSLFRKVLVYKTFGPFVLMSPFFKAHIDLMFSGPOLIKLIRVNDTAPD 360
QY 361 WQGFYTYVLLFVTAICLOTVLVHOVPHICFVSGMKRKTAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYTYVLLFVTAICLOTVLVHOVPHICFVSGMKRKTAVIGAVYRKALVITNSARKSTV 420
QY 361 WQGFYTYVLLFVTAICLOTVLVHOVPHICFVSGMKRKTAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYTYVLLFVTAICLOTVLVHOVPHICFVSGMKRKTAVIGAVYRKALVITNSARKSTV 420

QY 421 GEIYVLMSSVDAPRMDLATYINMTWGNPLQVYIALYILMTNLGSPVLGAVWVLMVPVN 480
DB 421 GEIYVLMSSVDAPRMDLATYINMTWGNPLQVYIALYILMTNLGSPVLGAVWVLMVPVN 480
QY 481 AVNAAKTKTYQVAAHMSKDKRIRKLMNEILNGIRYVLKLYAMELAFKDVLAIRBELVYLK 540
DB 481 AVNAAKTKTYQVAAHMSKDKRIRKLMNEILNGIRYVLKLYAMELAFKDVLAIRBELVYLK 540
QY 541 KSAVLSAVGFTWVCPFLVALCTFAVYVTIDENNIIIDAQTAFVSLAFNIRPLEPLILP 600
DB 541 KSAVLSAVGFTWVCPFLVALCTFAVYVTIDENNIIIDAQTAFVSLAFNIRPLEPLILP 600
QY 601 MVTSISIVQASVSLKRLRIFLSHELEPDSIERRPVKDGGGTSITVNAATFVARSPPPT 660
DB 601 MVTSISIVQASVSLKRLRIFLSHELEPDSIERRPVKDGGGTSITVNAATFVARSPPPT 660
QY 661 LMGITFSPGALVAVVGVCCKSSLSLSLALMBDVEGHVAKGSVAIVPQAMVIONT 720
DB 661 LMGITFSPGALVAVVGVCCKSSLSLSLALMBDVEGHVAKGSVAIVPQAMVIONT 720
QY 721 SLRENILFGCOLBEPYRSVIOACALLPDLIELPSGDRTEIGEGVNLSGGQKQVSLAR 780
DB 721 SLRENILFGQOLBEPYRSVIOACALLPDLIELPSGDRTEIGEGVNLSGGQKQVSLAR 780
QY 781 AVYSNADITYLPDPLSAVDHVGKHFENYIGPKMLKXKTRILVTHSMSTLPQVDVYIV 840
DB 781 AVYCDSDVYLLDPLSAVDHVGKHFENYIGPKMLKXKTRILVTHSMSTLPQVDVYIV 840
QY 841 MSGGKISSEMSYQELLARDGAFAPLRTYSTQEDQDAEENGVTGVSFGPKGAQKQNGM 900
DB 841 MSGGKISSEMSYQELLARDGAFAPLRTYSTQEDQDAEENGVTGVSFGPKGAQKQNGM 900
QY 901 LVTDLSAQLOLQOLSSSSYSYSGDISRHNSSTAELOKAEKEETWKLMEADKQTOGVKL 960
DB 901 LVTDLSAQLOLQOLSSSSYSYSGDISRHNSSTAELOKAEKEETWKLMEADKQTOGVKL 960
QY 961 SVYDWMKAIQGLFISLIFLPMCNHVSALASNYWLSMTDDPIVNGTOHTKRLSVYVG 1020
DB 961 SVYDWMKAIQGLFISLIFLPMCNHVSALASNYWLSMTDDPIVNGTOHTKRLSVYVG 1020
QY 1021 ALGISOGIAYFGYMAVSIIGIILASRCLHYDLHSILRSMPSPERTPSGTVNRFSEKL 1080
DB 1021 ALGISOGIAYFGYMAVSIIGIILASRCLHYDLHSILRSMPSPERTPSGTVNRFSEKL 1080
QY 1081 DTVDMSIPEVYIKMFMGSLFNVTGACIYIILATPIAIIIPPLGIYFFVORFYVASSROL 1140
DB 1081 DTVDMSIPEVYIKMFMGSLFNVTGACIYIILATPIAIIIPPLGIYFFVORFYVASSROL 1140
QY 1141 KRLESVSRSPIVSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENQAYPSIVANRWLA 1200
DB 1141 KRLESVSRSPIVSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENQAYPSIVANRWLA 1200
QY 1201 VRLCEVNCIVLPALFAVISRHSLSAGLVGLSVSYSLQVTTYVNLVRRSSEMETNIVA 1260
DB 1201 VRLCEVNCIVLPALFAVISRHSLSAGLVGLSVSYSLQVTTYVNLVRRSSEMETNIVA 1260
QY 1261 VERLKEYSFETKEAPMOIQDAPPSWPQYGRVFRNYCYRREDDLPFVLAHIVNTINGG 1320
DB 1261 VERLKEYSFETKEAPMOIQDAPPSWPQYGRVFRNYCYRREDDLPFVLAHIVNTINGG 1320
QY 1321 EKVGIAGRGAAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIAGRGAAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGSLRNMLDPPSQYSDDEWVTSLELAHLKQFVSLPKLDCHECAGEGNSVSGRQVYCL 1440
DB 1381 SGSLRNMLDPPSQYSDDEWVTSLELAHLKQFVSLPKLDCHECAGEGNSVSGRQVYCL 1440
QY 1441 ABALIRKTKILVDEAFAVDAVLETDLLQSTIRIQFQDDCTVLTAAHRLNTIMDTTRVYL 1500
DB 1441 ABALIRKTKILVDEAFAVDAVLETDLLQSTIRIQFQDDCTVLTAAHRLNTIMDTTRVYL 1500
QY 1501 DKGEIOBYGAPSDLLQORGLFYSMAXDAGLV 1531
DB 1501 DKGEIOBYGAPSDLLQORGLFYSMAXDAGLV 1531

QY 977 LSTFLEPCNHSALASYWMLSLMTDDPIVNGTQSHTKRLSVYCALGISOGIAVFGYSMA 1036
 Db 846 LSTFLEPCNHSALASYWMLSLMTDDPIVNGTQSHTKRLSVYCALGISOGIAVFGYSMA 905
 QY 1037 VSIIGGLIASHRCIHLHSLIRSPMSPEPSPCNLVNRSKEKIDPTVDSMTPEVTKPMFG 1096
 Db 906 VSIIGGLIASHRCIHLHSLIRSPMSPEPSPCNLVNRSKEKIDPTVDSMTPEVTKPMFG 965
 QY 1097 SLFNVIGACIVILLATPIAIIIPPLGLIYFVORFYVASSRQKRLSVSRSPVYSHFN 1156
 Db 966 SLFNVIGACIVILLATPIAIIIPPLGLIYFVORFYVASSRQKRLSVSRSPVYSHFN 1025
 QY 1157 ETLIGSVIRAFEEQERFIHOSDKVDENQKAYPSIYANMVLARLEVCNCLVLPAL 1216
 Db 1026 ETLIGSVIRAFEEQERFIHOSDKVDENQKAYPSIYANMVLARLEVCNCLVLPAL 1085
 QY 1217 FAVSRHSLSAGVGLSVSYSLQVTTYNMVLVRMSSEMETNIIVVERLKXSETEKAPW 1276
 Db 1086 FAVSRHSLSAGVGLSVSYSLQVTTYNMVLVRMSSEMETNIIVVERLKXSETEKAPW 1145
 QY 1277 QIOETAPSSWPQVGRVEFRNYCLRYREDLDFVLRIHNVITNGEKVIGVGTGAGKSSL 1336
 Db 1146 QIOETAPSSWPQVGRVEFRNYCLRYREDLDFVLRIHNVITNGEKVIGVGTGAGKSSL 1205
 QY 1337 TLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPDQVPLFSGSLRNNLDPSPQYSD 1396
 Db 1206 TLGLFRINESAEGEIIIDGINIAKIGLHDLRFKTTIIPDQVPLFSGSLRNNLDPSPQYSD 1265
 QY 1397 EEVWTSLEIHLKOPVSLPKLHCEGEGENISVQGRQVLCARALRLTKTLVLDEA 1456
 Db 1266 EEVWTSLEIHLKOPVSLPKLHCEGEGENISVQGRQVLCARALRLTKTLVLDEA 1385
 QY 1457 TAAVDLETDLDLQSTIRTOFEDCTVLTAHRLNTIMDTRYIVLDKGEIOEGYGAPSDLQ 1516
 Db 1326 TAAVDLETDLDLQSTIRTOFEDCTVLTAHRLNTIMDTRYIVLDKGEIOEGYGAPSDLQ 1385
 QY 1517 QRGIFYSMADAGLV 1531
 Db 1386 QRGIFYSMADAGLV 1400

RESULT 11
 MRP1_MOUSE STANDARD; PRT; 1528 AA.
 AC 035379;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Multidrug resistance-associated protein 1 (ATP-binding cassette sub-family C member 1).
 GN Name=Abcc1; Synonyms=Abcc1a, Abcc1b, Mdrap, Mrp;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; OC Muridae; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Muscle.
 RX MEDLINE=96251691; PubMed=8649356;
 RA Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P.C., Deeley R.G.;
 RT "Structure and expression of the messenger RNA encoding the murine multidrug resistance protein, an ATP-binding cassette transporter.";
 RL Mol. Pharmacol. 49:962-971(1996).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STEAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=22354683; PubMed=1246851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kanakawa T., Adachi J., Bono H., Kondo S., Nakai I., Osato N., Saito R., Suzuki H., Yamakata I., Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T., Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,

RA Schriul L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.B., Cousins S., Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S., Gasterland T., Gariboldi M., Giesi C., Godzik A., Gough J., RA Grimmond S., Gustich S., Hirokawa N., Jackson I.J., Jarvis E.D., Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L., RA Kanagaya A., Kurochkin I.V., Lee Y., Lemhard B., Lyons P.A., Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H., Nagaishi T., Numata K., Okada T., Pavan W.J., Pertez G., Pesole G., RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M., RA Savelin A., Schneider C., Sempé C.A., Setou M., Shimada K., Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M., RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C., RA Wilming L.G., Wysshaw-Boris A., Yanagisawa M., Yang I., Yang L., RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K., RA Shiraki T., Maki K., Kawai U., Aizawa K., Arikawa T., Fukuda S., Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I., RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A., RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 CC -1- FUNCTION: May participate directly in the active transport of drugs into subcellular organelles or influence drug distribution indirectly (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.
 CC -1- SIMILARITY: Contains 2 ABC transmembrane type-1 domains.
 CC -1- SIMILARITY: Contains 2 ABC transporter domains.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
 CC EMBL; AF022908; AAC80938.1; -; mRNA.
 CC EMBL; AK029876; AAC26654.1; -; mRNA.
 CC HSP, P08716; IMT0.
 CC Ensembl; ENSMUSG0000023088; Mus musculus.
 CC MG1; MG1; 102676; Abcc1.
 CC GO; GO:0005887; C:integral to plasma membrane; IDA.
 CC InterPro; IPR003593; AAA_ATPase.
 CC InterPro; IPR011527; ABC_membrane_1.
 CC InterPro; IPR001140; ABC_TM_transp.
 CC InterPro; IPR004399; ABC_transp_like.
 CC InterPro; IPR005292; MRP_assoc.
 CC Pfam; PF00664; ABC_membrane; 2.
 CC Pfam; PF00005; ABC_tran; 2.
 CC ProDom; PD000006; ABC_transporter; 2.
 CC SMART; SM00382; AAA_2.
 CC TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
 CC PROSITE; PS50929; ABC_TM1F; 2.
 CC PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
 CC PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
 CC ATP-binding; Glycoprotein; Nucleotide-binding; Repeat; Transmembrane; transport.
 KM TOPO_DOM 1 33 Extracellular (By similarity).
 FT TOPO_DOM 34 54 1 (By similarity).
 FT TOPO_DOM 55 74 Cytoplasmic (By similarity).
 FT TRANSMEM 75 95 2 (By similarity).
 FT TOPO_DOM 96 100 Extracellular (By similarity).
 FT TRANSMEM 101 121 3 (By similarity).
 FT TOPO_DOM 122 133 Cytoplasmic (By similarity).
 FT TRANSMEM 134 154 4 (By similarity).
 FT TOPO_DOM 155 172 Extracellular (By similarity).
 FT TRANSMEM 173 193 5 (By similarity).
 FT TOPO_DOM 194 317 Cytoplasmic (By similarity).
 FT TRANSMEM 318 338 6 (By similarity).
 FT TOPO_DOM 339 364 Extracellular (By similarity).

FT	TRANSNM	365	385	7 (By similarity).
FT	TOPO_DOM	386	441	Cytoplasmic (By similarity).
FT	TRANSNM	442	462	8 (By similarity).
FT	TOPO_DOM	463	465	Extracellular (By similarity).
FT	TRANSNM	466	486	9 (By similarity).
FT	TOPO_DOM	487	548	Cytoplasmic (By similarity).
FT	TRANSNM	549	569	10 (By similarity).
FT	TOPO_DOM	570	591	Extracellular (By similarity).
FT	TRANSNM	592	612	11 (By similarity).
FT	TOPO_DOM	613	663	Cytoplasmic (By similarity).
FT	TRANSNM	664	984	12 (By similarity).
FT	TOPO_DOM	985	1022	Extracellular (By similarity).
FT	TRANSNM	1023	1043	13 (By similarity).
FT	TOPO_DOM	1044	1086	Cytoplasmic (By similarity).
FT	TRANSNM	1087	1107	14 (By similarity).
FT	TOPO_DOM	1108	1108	Extracellular (By similarity).
FT	TRANSNM	1109	1129	15 (By similarity).
FT	TOPO_DOM	1130	1200	Cytoplasmic (By similarity).
FT	TRANSNM	1201	1221	16 (By similarity).
FT	TOPO_DOM	1222	1223	Extracellular (By similarity).
FT	TRANSNM	1224	1244	17 (By similarity).
FT	TOPO_DOM	1245	1528	Cytoplasmic (By similarity).
FT	DOMAIN	326	609	ABC transmembrane type-1.1.
FT	DOMAIN	644	868	ABC transporter 1.
FT	DOMAIN	971	1253	ABC transmembrane type-1.2.
FT	DOMAIN	1290	1524	ABC transporter 2.
FT	NP_BIND	678	685	ATP 1 (Potential).
FT	NP_BIND	1324	1331	ATP 2 (Potential).
FT	CARBOHYD	19	19	N-linked (GLCNAC. . .) (Potential).
FT	CARBOHYD	1003	1003	N-linked (GLCNAC. . .) (Potential).
SEQUENCE		1528 AA;	17185 MW;	68FD13667D61DBB CRC64;

Query Match 70.7%; Score 7002.5; DB 1; Length 1528; Best Local Similarity 88.0%; Pred. No. 0;

Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY	1	MALRGFCADGSDPLMDMNTMTNTPDKCFONTLVVWPCYTLACFPFYLYSRH	60
DB	1	MALRFSFADGSDPLMDMNTMTNTPDKCFONTLVVWPCYTLACFPFYLYSRH	60
QY	61	DRGYIQMTPLNKTKTALGFLMIVCMADLFYSFWERSRGIFLAPVFLVPTLLGITYLLA	120
DB	61	DRGYIQMTPLNKTKTALGFLMIVCMADLFYSFWERSRGIFLAPVFLVPTLLGITYLLA	120
QY	121	TFPIQLERRKGVSSGIMLTFMVALCALAIIKSKMTAKEDAOVDLFRDITFYFYS	180
DB	121	TFPIQLERRKGVSSGIMLTFMVALCALAIIKSKMTAKEDAOVDLFRDITFYFYS	180
QY	181	LILILQVLSCFSDRSPLESTIHDPNCPSSSASFLRITFWMTGLIVRGYRQPLEGSD	240
DB	181	LILILQVLSCFSDRSPLESTIHDPNCPSSSASFLRITFWMTGLIVRGYRQPLEGSD	240
QY	241	LMSLNKEDTSEOVVPLVKNWKECAKTRKQPVNVYS-SKDPAQPESSKVDANEVEA	299
DB	241	LMSLNKEDTSEOVVPLVKNWKECAKTRKQPVNVYS-SKDPAQPESSKVDANEVEA	299
QY	300	LIYKSPKXENNPSTLFKYLYTFPGPYFLMSFPFKAIHDLMPGSOIILKLIKPVNDTRAP	359
DB	301	LIYKSPKXENNPSTLFKYLYTFPGPYFLMSFPFKAIHDLMPGSOIILKLIKPVNDTRAP	359
QY	360	DMOGYFTVLLFYTAGCQTLVHOYFHCIVSGKRITAVTGVYRRLVITNSARST	419
DB	361	DMOGYFTVLLFYTAGCQTLVHOYFHCIVSGKRITAVTGVYRRLVITNSARST	419
QY	420	VGEIVNLMSYDAGFMDLATYINMWSAPLOVILATYLMNLGSPVLAGAVAVLVAVPV	479
DB	421	VGEIVNLMSYDAGFMDLATYINMWSAPLOVILATYLMNLGSPVLAGAVAVLVAVPV	479
QY	480	NAYWAKTKTYOYVAHMSKNRITKLMEIINGRIYVCLKYAMELAFKDKVLAIROBELKVL	539
DB	481	NAYWAKTKTYOYVAHMSKNRITKLMEIINGRIYVCLKYAMELAFKDKVLSIROBELKVL	540
QY	540	KKSAYLSAVGTFTWVCTPFLVALCTFAVYVYTIIDENNILDAQTLAFVSLAFNIIARFPLNLT	599

DB	541	KKSAYLSAVGTFTWVCTPFLVALCTFAVYVYTIIDENNILDAQTLAFVSLAFNIIARFPLNLT	600
QY	600	PWYISSIVQASVSLKRLRITLSHELEPDSIERRPVQGGGINSITVRATFTTAABDDP	659
DB	601	PWYISSIVQASVSLKRLRITLSHELEPDSIERRSIKSGEG-NSITVKNAFTTAAABDDP	659
QY	660	TLNGITPSIEGALVAVGVGGKSSLSALAEKMDKVGHAIVKSVVYVPOQAION	719
DB	660	TLNGITPSIEGALVAVGVGGKSSLSALAEKMDKVGHAIVKSVVYVPOQAION	719
QY	720	DSLRENTLFCQLEEPYRSVIOACALLPDLLEILPSGDRTEIGEGVNLGGQKQVSLA	779
DB	720	DSLRENTLFCQLEEPYRSVIOACALLPDLLEILPSGDRTEIGEGVNLGGQKQVSLA	779
QY	780	RAYSNADYLPDDPLSAVDAAVGHKIFENVIGPKMLKOKTRILVTHSMYLPQVDYII	839
DB	780	RAYSNADYLPDDPLSAVDAAVGHKIFENVIGPKMLKOKTRILVTHSMYLPQVDYII	839
QY	840	VMSGKISEMGSYOELLARDAFAELRTASTVEQEDAEENGVTGVSGRGKAKOMENG	899
DB	840	VMSGKISEMGSYOELLARDAFAELRTASTVEQEDAEENGVTGVSGRGKAKOMENG	899
QY	900	MLVDSAGKQLOKOLSSSSSYSGDISRHHNSTABLQKAEKKEETWKLMEADKRAQTQVY	959
DB	900	MLVDSAGKQLOKOLSSSSSYSGDISRHHNSTABLQKAEKKEETWKLMEADKRAQTQVY	959
QY	960	LSYVMDYMKALIGLIFSLIFLFCMCHVSAASNYMLSTWTD-PIVNGTOEHTKRLSV	1018
DB	960	LSYVMDYMKALIGLIFSLIFLFCMCHVSAASNYMLSTWTD-PIVNGTOEHTKRLSV	1018
QY	1019	YGALGISQGIYAVGYSNAVSIIGLILASRCHLVLLHSILASPMSPFERTSSGNLVNRFK	1078
DB	1019	YGALGISQGIYAVGYSNAVSIIGLILASRCHLVLLHSILASPMSPFERTSSGNLVNRFK	1078
QY	1079	ELDTVDSMPEVYKMFGLFENYIGACIVTLLATPAAIIIPPLGLYFPFYQRYVASSR	1138
DB	1079	ELDTVDSMPEVYKMFGLFENYIGACIVTLLATPAAIIIPPLGLYFPFYQRYVASSR	1138
QY	1136	QLKRLBSVSRSPYSHNETLLGVSVYRAFEQBERFIHQSDLKVDENQKAYPSIVANRW	1195
DB	1136	QLKRLBSVSRSPYSHNETLLGVSVYRAFEQBERFIHQSDLKVDENQKAYPSIVANRW	1195
QY	1199	LAVRLCEVNCIYLPALPAVISRHSLSAGLVGLSVYSIQVTTYLNLVYRMSSEMTNI	1258
DB	1199	LAVRLCEVNCIYLPALPAVISRHSLSAGLVGLSVYSIQVTTYLNLVYRMSSEMTNI	1258
QY	1259	VAVERLKEYSETEKEAPMOIQETAPPSWQVGRVREPRNYCLARYREDLDFVLRIHINTIN	1318
DB	1259	VAVERLKEYSETEKEAPMOIQETAPPSWQVGRVREPRNYCLARYREDLDFVLRIHINTIN	1318
QY	1319	GGEKVGIVGTGAKSLLTGLFRINESAGEIILIDGINIAKTGLDLRFRKTIIPDDV	1378
DB	1319	GGEKVGIVGTGAKSLLTGLFRINESAGEIILIDGINIAKTGLDLRFRKTIIPDDV	1378
QY	1379	LFGSGLRMLNDPPSOYSDEEVTSLBLAHKDFVSALPDKLDEHCBGGENLSVGOQOLV	1438
DB	1379	LFGSGLRMLNDPPSOYSDEEVTSLBLAHKDFVSALPDKLDEHCBGGENLSVGOQOLV	1438
QY	1439	CLARALLRKTIVLEATAVVDLETDLLOSTIRFOPEBCYTLTAHRLNTIMDTYRVI	1498
DB	1439	CLARALLRKTIVLEATAVVDLETDLLOSTIRFOPEBCYTLTAHRLNTIMDTYRVI	1498
QY	1499	VLDKGEIOEYGAPSDLLQGRGLFYSAKDAGLV	1531
DB	1499	VLDKGEIOEYGAPSDLLQGRGLFYSAKDAGLV	1531

RESULT 12
0810E4_RAT PRELIMINARY; PRT; 1532 AA.
ID 0810E4_RAT
AC 0810E4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
DE ATP-binding cassette protein C1.
GN Name=Abcc1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley, TISSUE=Spleen.
RA Yabuchi H., Takayanagi S., Ishikawa T.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -; SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AF487549; AAC05437.1; -; mRNA.
DR HSSP; P08716; IMTO.
DR Ensembl; ENSRNOG0000032748; Rattus norvegicus.
DR RGD; 3112; Abcc1.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0006106; F:nucleotide binding; IEA.
DR GO; GO:000610; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR001395; Aldol_ket_red.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA_2.
DR TIGRfams; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS0929; ABC_TWIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00063; ALDO-KETO_REDUCTASE_3; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1532 AA; 171493 MW; 2E639F63F5A3F68 CRC64;

Query Match 70.0%; Score 6932.5; DB 2; Length 1532;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 1339; Conservative 99; Mismatches 92; Indels 3; Gaps 3;

QY 1 MALRGFSADGSDPLMDMNTWNTSNPDFTKCFQNTVLVWPCRYLMACFPFYLYLSRH 60
DB 1 MALSFCSDSGSDPLMDMNTWNTSNPDFTKCFQNTVLVWPCRYLMACFPFYLYLSRH 60
QY 61 DRGIQMTPLNKKTALGFILMVICWADLFYSFWERSRGIFLAVFLVSPILLGITLLA 120
DB 61 DRGIQMTPLNKKTALGFILMVICWADLFYSFWERSRGIFLAVFLVSPILLGITLLA 120
QY 121 TFLIQLERRKVGSSGIMLTFMVALVCALALILRSKIMTALKEDAQVDLFFDITFYVFS 180
DB 121 TFLIQLERRKVGSSGIMLTFMVALVCALALILRSKIMTALKEDAQVDLFFDITFYVFS 180
QY 181 LLLIQLVLSCPSDRSPLESETIHDNPPCPESSASFLSRITTWITGLIVRGYROPLEGSD 240
DB 181 LLLIQLVLSCPSDRSPLESETIHDNPPCPESSASFLSRITTWITGLIVRGYROPLEGSD 240
QY 241 LMSLNKEDTSQGVVPLVYKMKKCECAKTRKOPVAVVYS-SKDPNQPKSSKVDANEVEA 299
DB 241 LMSLNKEDTSQGVVPLVYKMKKCECAKTRKOPVAVVYS-SKDPNQPKSSKVDANEVEA 299
QY 300 LIVSPKEMWNPSTFKVLYKTFEGPYFLMSFFKAIHDLMPGSPQILKLLKFPVNDTAP 359
DB 300 LIVSSHKDRPSTFKVLYKTFEGPYFLMSFFKAIHDLMPGSPQILKLLKFPVNDTAP 359
QY 360 DMGQYFTVLLFVACIQTVLVHOYFHIQVSGMKIKTAVIGAVYRKALVITNSARKST 419
DB 360 DMGQYFTVLLFVACIQTVLVHOYFHIQVSGMKIKTAVIGAVYRKALVITNSARKST 419

DB 361 DMGQYFTVLLFVACIQTVLVHOYFHIQVSGMKIKTAVIGAVYRKALVITNSARKST 420
QY 420 VGEIVNMSVDAQRFMDLATYINNMWSAPLOVIALYLWLNIGSPYLAGVAVVWLVMPV 479
DB 421 VGEIVNMSVDAQRFMDLATYINNMWSAPLOVIALYLWLNIGSPYLAGVAVVWLVMPV 480
QY 480 NAWMAKTKTYQVAHMSKDNRIKANEIINGIKVLKLYAMELAFKQKVLAIROEELKVL 539
DB 481 NAWMAKTKTYQVAHMSKDNRIKANEIINGIKVLKLYAMELAFKQKVLAIROEELKVL 540
QY 540 KKSATVSAVGFPMVCPPLVALCTFAVYVTIDENNILUDQOTAFVSLAPFNLFPNL 599
DB 541 KKSATVLAAGVFTWCTPFLALSTFAVYVDEKNILDKAKKAVSLAPFNLFPNL 600
QY 600 PMVSSIVQSVSLKRIKRIFLSHELEPDSIERPVDGSGTNSITVRNAFTWASDDP 659
DB 601 PMVSSIVQSVSLKRIKRIFLSHELEPDSIERPVDGSGTNSITVRNAFTWASDDP 660
QY 660 TLNGITSPGALVAVVGVGCGKSLLSALLAEMDKVEGHVAIKOSVAVVPOQAMION 719
DB 661 TLNGITSPGALVAVVGVGCGKSLLSALLAEMDKVEGHVAIKOSVAVVPOQAMION 720
QY 720 DSIKENILPQCQLEEPYRSVIACALLPLEILPSGDRTEIGKGVNLGGQKQKRSIA 779
DB 721 DSIKENILPQCQLEEPYRSVIACALLPLEILPSGDRTEIGKGVNLGGQKQKRSIA 780
QY 780 RAVYSNADIVLFDPLSAVDAAHVKHIFENVIGPKGMKNKTRILVTHSMSYLPQVDVII 839
DB 781 RAVYCNSDIYLDPLPSAVDAHVKHIFENVIGPKGMKNKTRILVTHSMSYLPQVDVII 840
QY 840 VMSGKISEMGSYOELLARDGAPAEPLRTYASTFEQEDABENGVTGVSGPEKAKQWENG 899
DB 841 VMSGKISEMGSYOELLARDGAPAEPLRTYASTFEQEDABENGVTGVSGPEKAKQWENG 900
QY 900 MLVYDSAGKQQLRQLSSSSSGDISRHNSSTELQKAEKKEBTMLMADKQKQYK 959
DB 901 ILVYDAVGKPLQRLSSSSSVVTVNQOHSVTELQKS-GYKEBTMLMADKQKQYK 959
QY 960 LSVYWDPMKAIGLPISTLSIFLFWCMNVASALASNYWLSLWTD-DPIVNGTOEHTKVL 1018
DB 960 LSVYWDPMKAIGLPISTLSIFLFWCMNVASALASNYWLSLWTD-DPIVNGTOEHTKVL 1019
QY 1019 YGALGISQGIAPFGVSAVSIIGILASRCLHVDLHLSILRSPMSFFERTSGNIVNPSK 1078
DB 1020 YGALGISQGIAPFGVSAVSIIGILASRCLHVDLHLSILRSPMSFFERTSGNIVNPSK 1079
QY 1079 ELDTVDSDMIPVYIKMFNGSLFNVIAGACTIVLLATPIAAIIIPPLGLTYFPQRYVASSR 1138
DB 1080 ELDTVDSDMIPVYIKMFNGSLFNVIAGACTIVLLATPIAAIIIPPLGLTYFPQRYVASSR 1139
QY 1139 QLRLESVSRSPPVYSHNETLLGVSIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRW 1198
DB 1140 QLRLESVSRSPPVYSHNETLLGVSIRAFEEQERFIHQSDLKVDENQKAYYPSIVANRW 1199
QY 1199 LAVALCEVCNCTVLAALFAVVISRHSLSAGLVGLSVSYSLQVTTYLMVLRMSSEMETNI 1258
DB 1200 LAVALCEVCNCTVLAALFAVVISRHSLSAGLVGLSVSYSLQVTTYLMVLRMSSEMETNI 1259
QY 1259 VAVERLKEYSTEEKAAWQIOETAPBSSWPQGVVERRANCLARREDLPVLRHINVTIN 1318
DB 1260 VAVERLKEYSTEEKAAWQIOETAPBSSWPQGVVERRANCLARREDLPVLRHINVTIN 1319
QY 1319 GGEKVGVGRTRGAGKSLTGLPFINSGAGEIITDGINIAKGLHDLRFITIIIPDDPV 1378
DB 1320 GGEKVGVGRTRGAGKSLTGLPFINSGAGEIITDGINIAKGLHDLRFITIIIPDDPV 1379
QY 1379 LFGSGLRMLNDLPFSQYSDDEEWTSLBLAHLKDFFVSALPDKLDHECAEGENTLVGQROLV 1438
DB 1380 LFGSGLRMLNDLPFSQYSDDEEWTSLBLAHLKDFFVSALPDKLDHECAEGENTLVGQROLV 1439
QY 1439 CLARALLRKTKILVLDAVTAADVLETDLLIQSTIRTOFEDSTVLTIAHRLNTIMDYRVI 1498
DB 1440 CLARALLRKTKILVLDAVTAADVLETDLLIQSTIRTOFEDSTVLTIAHRLNTIMDYRVI 1499

QY 1319 GGEKVGIVGRTGACKSSITLGLFRINSAEGEIIIDGINIAKIGLHDIRFKITIIIPDDPV 1378
DB 1320 GGEKVGIVGRTGACKSSITLGLFRINSAEGEIIIDGINIAKIGLHDIRFKITIIIPDDPV 1379
QY 1379 LFGSSILMNLDPEGOYSDPEFWTSLAHLKDFYSALPDKLDHRCAGGEMLSVGQRLV 1438
DB 1380 LFPSSLRNLDLPFQOYSDPEFWMALEHLKGFVSALPDKLNHECAGGEMLSVGQRLV 1439
QY 1439 CLARALRKTILVLDENATAVDELTDLIQSTRTPQEDCTVLITIAHLNTIMDYTRVI 1498
DB 1440 CLARALRKTILVLDENATAVDELTDLIQSTRTPQEDCTVLITIAHLNTIMDYTRVI 1499
QY 1499 VLDKGEIOEGAPSDLLQQRGLFYSMAKAGLV 1531
DB 1500 VLDKGEIRECGAPSELLQQRGVFYSMAKAGLV 1532
RESULT 14
Q810G9_RAT PRELIMINARY; PRT; 1523 AA.
ID Q810G9_RAT PRELIMINARY; PRT; 1523 AA.
AC Q810G9;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE ATP-binding cassette protein C1 variant A.
GN Name=Abcc1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Spleen;
RA Takayanagi S., Ichikawa T.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY174892; AAC44983.1; -, mRNA.
DR HSSP; P08716; IMT0.
DR RGD; 3112; Abcc1.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR001395; Aldo/ket_red.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS00929; ABC_TMIF; 2.
DR PROSITE; PS00811; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00063; ALDOXETO_REDUCTASE_3; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
KW SEQUENCE 1523 AA; 170505 MW; E40337051A1CB9C6 CRC64;
Query Match 69.6%; Score 6892; DB 2; Length 1523;
Best Local Similarity 87.0%; Pred. No. 0;
Matches 1333; Conservative 98; Mismatches 90; Indels 12; Gaps 4;
QY 1 MALRFGCAGDSDDPLMDNMVNTSNPPTKCFONTLVWVPCFYLMACFPFYFLYLSRH 60
DB 1 MALRFGCAGDSDDPLMDNMVNTSNPPTKCFONTLVWVPCFYLMACFPFYFLYLSRH 60
QY 61 DRGYIOMTPLNKXTALGFLLMIWCMADLFYSFWERSRGIPLAPVFLVSPILLGITTLA 120

DB 61 DRGYIOMTHLNKATLALGFLLMIWCMADLFYSFWERSRQGMILAVLLVSPILLGITTLA 120
QY 121 TFLIOLERRKGVSSGIMLFWLVAVALCALAIIIRSKIMTALKEDAQVDLFPDITFYVFS 180
DB 121 TFLIOLERRKGVSSGIMLFWLVAVALCALAIIIRSKIIISALKKDAQVDMFDSAFYLFT 180
QY 181 LLLIQLVLSGFSRDRPLFSETIHDNPNCPSSSASFLSRIITFWITGLIWRGQRLPBGSD 240
DB 181 LVFIQLVLSCPSDSSPLFSEIVRDPNCPSSSASFLSRIITFWITGLIWRGQRLPBGSD 240
QY 241 LMSLNKEDTSEQVPLVVKWKKECATRKQPKVYVS-SKDPQPKSSAVDANEVEA 299
DB 241 LMSLNKEDTSEQVPLVVKWKKECATRKQPKVYVS-SKDPQPKSSAVDANEVEA 299
QY 300 LIVESPOKEMNPSLFKVLYKTFGPFYFMSFFKAIHDLMSPGQIILKIFVNDTKAP 359
DB 301 LIVKSHKDRDPSLFKVLKTFGPFYFMSFFKAIHDLMSPGQIILKIFVNDTKAP 359
QY 360 DMQGYFYTVLLFVTACIQLTLVHQYFHCIVSGNRITKAVIAGAVYKALVITNSARKST 419
DB 361 DMQGYLYTALLFVACIQLTLVHQYFHCIVSGNRITKAVIAGAVYKALVITNSARKST 420
QY 420 VGEIYVNLMSVDAQFPMDLATYINMINSAPLOVILALVLMNLGSPSVLAGVAVWLVVPV 479
DB 421 VGEIYVNLMSVDAQFPMDLATYINMINSAPLOVILALVLMNLGSPSVLAGVAVWLVVPV 480
QY 480 NAWAMKTKYQVAMHMSKDNRIKLMBEILNGIYVLYLWAMELAFKQKVLAIROBELKVL 539
DB 481 NAWAMKTKYQVAMHMSKDNRIKLMBEILNGIYVLYLWAMELAFKQKVMIRBELKVL 540
QY 540 KKSAYLSAVGTFTWCTPPLVALCTPAVYTTIDENILDAQTAVALFNILRPLNLT 599
DB 541 KKSAYLSAVGTFTWCTPPLVALCTPAVYTTIDENILDAQTAVALFNILRPLNLT 600
QY 600 PMVTSIYQASVSLKRLRIFLSHEELBPDSEIRPVDGCGTNSITTRNATFTAKRSPP 659
DB 601 PMVTSIYQASVSLKRLRIFLSHEELBPDSEIRSIKDGGMNSIYVNAFTFARDEPP 660
QY 660 TLNGITTSIPGALVAVVGVGCGKSSLSLALMBDKVEGVAKSVAVVPOQAMION 719
DB 661 TLNGITTSIPGALVAVVGVGCGKSSLSLALMBDKVEGVAKSVAVVPOQAMION 720
QY 720 DLSRENILFGCOLPEPYRYVQACALLPDLEILPSGDREIGKGVNLGGQKORVSLA 779
DB 721 DLSRENILFGCOLPEPYRYVQACALLPDLEILPSGDREIGKGVNLGGQKORVSLA 780
QY 780 RAVYSNADITYLPDPLSAVDAAVGHKIFENVYIGKGMKAKTRILVTHSMYLPQVDYII 839
DB 781 RAVYCNSDIYLLDPLSAVDAAVGHKIFENVYIGKGMKAKTRILVTHSMYLPQVDYII 840
QY 840 VMSGKTSSEMGSYOELLARDGAFAPLRTYASTBOEADAEENGYGVSGPKAKOMENG 899
DB 841 VMSGKTSSEMGSYOELLARDGAFAPLRTYASTBOEADAEENGYGVSGPKAKOMENG 900
QY 900 MLVYDSAGKOLQROLSSSSYSYSDISRHNSHTAEILOKAEAKKETWKLMEADKAQYGVK 959
DB 901 MLVYDSAGKOLQROLSSSSYSYSDISRHNSHTAEILOKAEAKKETWKLMEADKAQYGVK 959
QY 960 LSVYWDYKAIQLFISPLSTFLPMONHVSALASNTWLSLWTD-PIYNGTOEHKVRLSV 1018
DB 961 LSVYWDYKAIQLFISPLSTFLPMONHVSALASNTWLSLWTD-PIYNGTOEHKVRLSV 1018
QY 951 LSVYWNWYKAIQLFISPLSTFLPMONHVSALASNTWLSLWTD-PIYNGTOEHKVRLSV 1010
DB 951 LSVYWNWYKAIQLFISPLSTFLPMONHVSALASNTWLSLWTD-PIYNGTOEHKVRLSV 1010
QY 1019 YGALIGISGIVAPVSNMVSIGLILASCHVDLILSLRSPMSFFERTPSGNTVNRPSK 1078
DB 1011 YGALIGISGIVAPVSNMVSIGLILASCHVDLILSLRSPMSFFERTPSGNTVNRPSK 1070
QY 1079 ELDTVDSMIPVIVIKFMGSLFNVIGACIVILATPAAIIIPPLGLIYFYQRFVASSR 1138
DB 1071 ELDTVDSMIPVIVIKFMGSLFNVIGACIVILATPAAIIIPPLGLIYFYQRFVASSR 1130
QY 1139 QKRLSEVSRSPPVSHFNETLLGVSVIRAFEEQERFIHQSLKVDENOKAYYPSIVANRW 1198

Db 1131 QKRLSESVSRSPVSHFNETLLGVSVIRAFEEORERFIKQSDLKVDENQKAYPSIVANRW 1190
QY 1199 LAVERLECGNCIVT.PAALFAVISRHSLSAGVGLSVSYSLQVTTYIMLVMSSEMENI 1258
Db 1191 LAVERLECGNCIVT.PAALFAVISRHSLSAGVGLSVSYSLQVTTYIMLVMSSEMENI 1250
QY 1259 VAVERLEKSESTKEAPWQIOETAPSSSWPQVGRFEPFNCLARYREDLDFLRHINVTIN 1318
Db 1251 VAVERLEKSESTKEASWQIOETAPSPWPHSGRVERPFDYCLARKREDLVLKINVTIE 1310
QY 1319 GGEVGVIGRTGAGKSSITLGLFRINSEAGEEIIIDGINAKIGLHDLRFKITTIPDPV 1378
Db 1311 GGEVGVIGRTGAGKSSITLGLFRINSEAGEEIIIDGINAKIGLHDLRFKITTIPDPV 1370
QY 1379 LFGSGLSMNLDPEGOYSDPEEWTSLELAHLKDFPSALPDKLDHRCAGEGENLSTVGORLV 1438
Db 1371 LFGSGLSMNLDPEGOYSDPEEWTSLELAHLKDFPSALPDKLDHRCAGEGENLSTVGORLV 1430
QY 1439 CLAAALRKTKILVLDENTAAVDETDLIOSTIRTOPEDCVTLTIAHRLNTIMDYTRVI 1498
Db 1431 CLAAALRKTKILVLDENTAAVDETDLIOSTIRTOPEDCVTLTIAHRLNTIMDYTRVI 1490
QY 1499 VLDKGEIOEYGAPSDLLQORGLFYSMAKAGLV 1531
Db 1491 VLDKGEIOEYGAPSDLLQORGLFYSMAKAGLV 1523

RESULT 15
QSF364_CHICK
ID QSF364_CHICK PRELIMINARY; PRT; 1525 AA.
AC QSF364;
DT 10-MAY-2005 (Tremblrel. 30, Created)
DT 10-MAY-2005 (Tremblrel. 30, Last sequence update)
DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=RCJMB04.32d20;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CH: TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezubov Y., Zaim J.,
RA Plachy J., Carnuci P., Blasodatski A., Kostovska D., Kotler M.,
RT "Full-length cDNAs from chicken bursa lymphocytes to facilitate
RT genefunction analysis";
RL Genome Biol. 6:R6(2005).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AJ851786; CAH65420.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR Prodom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PSS0929; ABC_TMIF; 2.
DR PROSITE; PSS0211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PSS0893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PSS0107; PROTEIN_KINASE_ATP; UNKNOWN_1.

KW ATP-binding; Hypothetical protein; Nucleotide-binding; Repeat;
KW Transport. 1525 AA; 170972 MW; 3872486711B878FD CR664;
SQ SEQUENCE 1525 AA; 170972 MW; 3872486711B878FD CR664;
Query Match 63.6%; Score 6301; DB 2; Length 1525;
Best local similarity 77.6%; Pred. No. 0;
Matches 1199; Conservative 176; Mismatches 148; Indels 20; Gaps 7;
QY 1 MALRGFSSAGSDPLMNMNTMTNTSNDPFCQNTLVWVPCYTLVWCPFFYLYSRH 60
Db 1 MGIESLSAASAEFFWMNLTWHTENDPFOCFQNTLVWVPCYTLVWCPFFYLYSRH 60
QY 61 DRGYIQMTPLNKTALGFLIMTYCMADLFPSFERSRGIFLAFVPLVSPFLGITTLLA 120
Db 61 DRGYIQMTPLNKTALGFLIMTYCMADLFPSFERSRGIFLAFVPLVSPFLGITTLLA 120
QY 121 TFLIQHRLKGVQSSGVMTLFWLISLCAVIFPSKIMLALNTDEVDARFVYFCYFI 180
Db 121 TFLIQHRLKGVQSSGVMTLFWLISLCAVIFPSKIMLALNTDEVDARFVYFCYFI 180
QY 181 LLLIQVLVSCPSDRSLFSETIHDNPNCPSASFLSRITFWWTGLIVGYPOLGSD 240
Db 181 LLLIQVLVSCPSDRSLFSETIHDNPNCPSASFLSRITFWWTGLIVGYPOLGSD 240
QY 181 LLLVQLILSCFPEKPLFSEAVNDPKCPFSASFLSRITFWWTGLIMIQHRRPLAKD 240
Db 181 LLLVQLILSCFPEKPLFSEAVNDPKCPFSASFLSRITFWWTGLIMIQHRRPLAKD 240
QY 241 LMSLNKEDTSBOYVPLVKNMKKECAKTRKQPVKVVSSDPAQPKESKVDAN----EE 296
Db 241 LMSLNKEDTSBOYVPLVKNMKKECAKTRKQPVKVVSSDPAQPKESKVDAN----EE 296
QY 297 VEALIVKSPQKWNPSLFKVLVYKTFGPFLMSFFPKAIDHLMFSGPQILKLIKPVNDT 356
Db 297 VEALIVKSPQKWNPSLFKVLVYKTFGPFLMSFFPKAIDHLMFSGPQILKLIKPVNDT 356
QY 296 ABLIITKPSGRSSASLSTVLYKTFPGPYFLMSLFKKAHDLMFTGEIILKILINPVNK 355
Db 296 ABLIITKPSGRSSASLSTVLYKTFPGPYFLMSLFKKAHDLMFTGEIILKILINPVNK 355
QY 357 KAPDMQGYFTVLLVFTACIQTLYLHOYFHIQVYSGKRITAVIQAAYKALVITNSARK 416
Db 357 KAPDMQGYFTVLLVFTACIQTLYLHOYFHIQVYSGKRITAVIQAAYKALVITNSARK 416
QY 356 SAPWQGYFTVGLLFWCACIQTLLHQYFHIQVYSGKRITAVIQAAYKALVITNSARK 415
Db 356 SAPWQGYFTVGLLFWCACIQTLLHQYFHIQVYSGKRITAVIQAAYKALVITNSARK 415
QY 417 SSTVGEIVINMSVDAQEFMDLAVYINMISAPLOVILATLVLMNIGPSVLGAVNWM 476
Db 417 SSTVGEIVINMSVDAQEFMDLAVYINMISAPLOVILATLVLMNIGPSVLGAVNWM 476
QY 416 TSTVGEIVINMSVDAQEFMDLAVYINMISAPLOVILATLVLMNIGPSVLGAVNWM 475
Db 416 TSTVGEIVINMSVDAQEFMDLAVYINMISAPLOVILATLVLMNIGPSVLGAVNWM 475
QY 477 VPVNAVAMAKTKTYQVAHMKSKONRIKLMNEIINGIKVLKLYAMELAFKQVLAIROBEL 536
Db 477 VPVNAVAMAKTKTYQVAHMKSKONRIKLMNEIINGIKVLKLYAMELAFKQVLAIROBEL 536
QY 476 VPINAVAMAKTKTYQVAHMKSKONRIKLMNEIINGIKVLKLYAMELAFKQVLAIROBEL 535
Db 476 VPINAVAMAKTKTYQVAHMKSKONRIKLMNEIINGIKVLKLYAMELAFKQVLAIROBEL 535
QY 537 KVLKKSAYLSAVGTFPTVCTPFLVALCTPAVYVYTTIDENNITLDAQAFVSLALFNILRPPL 596
Db 537 KVLKKSAYLSAVGTFPTVCTPFLVALCTPAVYVYTTIDENNITLDAQAFVSLALFNILRPPL 596
QY 536 KVLKKSAYLSAVGTFPTVCTPFLVALCTPAVYVYTTIDENNITLDAQAFVSLALFNILRPPL 595
Db 536 KVLKKSAYLSAVGTFPTVCTPFLVALCTPAVYVYTTIDENNITLDAQAFVSLALFNILRPPL 595
QY 597 NILPMYSSIVQASVSLKRLRFLSHEELPDSIERPVQDGGTNSITVRNATFTPARS 656
Db 597 NILPMYSSIVQASVSLKRLRFLSHEELPDSIERPVQDGGTNSITVRNATFTPARS 656
QY 596 NILPMYSSIVQASVSLKRLRFLSHEELPDSIERPVQDGGTNSITVRNATFTPARS 653
Db 596 NILPMYSSIVQASVSLKRLRFLSHEELPDSIERPVQDGGTNSITVRNATFTPARS 653
QY 657 DPTLNGITPISIEGALVAVVGVGCGSKSLALAEQDKVGHVAKISVAVVPOQAM 716
Db 657 DPTLNGITPISIEGALVAVVGVGCGSKSLALAEQDKVGHVAKISVAVVPOQAM 716
QY 654 DPTSLNSINTVPEGSLIYAVGVGCGSKSLALAEQDKVGHVAKISVAVVPOQAM 713
Db 654 DPTSLNSINTVPEGSLIYAVGVGCGSKSLALAEQDKVGHVAKISVAVVPOQAM 713
QY 717 IQNDSIRENILLFSCQLEEPYRSVIOACALLPDLLEILPSGDRTEIGEGVNLGGQKQV 776
Db 717 IQNDSIRENILLFSCQLEEPYRSVIOACALLPDLLEILPSGDRTEIGEGVNLGGQKQV 776
QY 777 SLARAYSNADITLFPDPLSAVNAHVKHIFENVIGPKMLKTKTRLVTHSHSYLPQVD 836
Db 777 SLARAYSNADITLFPDPLSAVNAHVKHIFENVIGPKMLKTKTRLVTHSHSYLPQVD 836
QY 774 SLARAYCNADITLFPDPLSAVNAHVKHIFENVIGPKMLKTKTRLVTHSHSYLPQVD 833
Db 774 SLARAYCNADITLFPDPLSAVNAHVKHIFENVIGPKMLKTKTRLVTHSHSYLPQVD 833
QY 837 VITVSGKISIEGVSQELIARDAFAEFLRTASTBOEDARENGTVGSGP-GKEAKQ 895
Db 837 VITVSGKISIEGVSQELIARDAFAEFLRTASTBOEDARENGTVGSGP-GKEAKQ 895
QY 834 TILVMTDEISEMGVSQELIKDGAFAEFLRTASTBOEDARENGTVGSGP-GKEAKQ 888
Db 834 TILVMTDEISEMGVSQELIKDGAFAEFLRTASTBOEDARENGTVGSGP-GKEAKQ 888
QY 896 MEGMLVTSAGAQLOLQOLSSSSSYGDI--SHHNSIAFLQCAEAKSETWTLMADKA 953
Db 896 MEGMLVTSAGAQLOLQOLSSSSSYGDI--SHHNSIAFLQCAEAKSETWTLMADKA 953
QY 889 VENGVLVNDATGKLMRQSLNSSTYRETKGSHQSTALQKPLAK-NSMKTLEADTA 947
Db 889 VENGVLVNDATGKLMRQSLNSSTYRETKGSHQSTALQKPLAK-NSMKTLEADTA 947
QY 954 QTVQVKLIVYVMDYMKAIIGLFIISFLIFLFCNHNVSALASVYMLSLWTDPIVNGTQEHK 1013
Db 954 QTVQVKLIVYVMDYMKAIIGLFIISFLIFLFCNHNVSALASVYMLSLWTDPIVNGTQEHK 1013

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 15:05:25 ; Search time 39.4867 Seconds
(without alignments)
4076.547 Million cell updates/sec

Title: US-10-665-283-8

Perfect score: 10016

Sequence: 1 MALRGCASAGSDPLMDMNV.....RSVAVAKKPKFSIPDSLS 1947

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/1/1aa/5 COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/6 COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/H COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/PCTUS COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/RB COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7860	78.5	1531	1 US-08-463-092B-4	Sequence 4, Appl1
2	7860	78.5	1531	1 US-08-462-109A-4	Sequence 4, Appl1
3	7860	78.5	1531	1 US-08-460-907B-4	Sequence 4, Appl1
4	7860	78.5	1531	2 US-08-463-179A-4	Sequence 4, Appl1
5	7860	78.5	1531	2 US-08-461-384B-4	Sequence 4, Appl1
6	7860	78.5	1531	2 US-09-647-140B-19	Sequence 19, Appl1
7	7849	78.4	1531	1 US-08-141-893-2	Sequence 2, Appl1
8	7849	78.4	1531	1 US-08-463-092B-2	Sequence 2, Appl1
9	7849	78.4	1531	1 US-08-462-109A-2	Sequence 2, Appl1
10	7849	78.4	1531	1 US-08-460-907B-2	Sequence 2, Appl1
11	7849	78.4	1531	2 US-08-463-179A-2	Sequence 2, Appl1
12	7849	78.4	1531	2 US-08-461-384B-2	Sequence 2, Appl1
13	7849	78.4	1531	2 US-08-407-207A-2	Sequence 2, Appl1
14	7002.5	69.9	1528	1 US-08-463-092B-6	Sequence 6, Appl1
15	7002.5	69.9	1528	1 US-08-462-109A-6	Sequence 6, Appl1
16	7002.5	69.9	1528	1 US-08-460-907B-6	Sequence 6, Appl1
17	7002.5	69.9	1528	2 US-08-463-179A-6	Sequence 6, Appl1
18	7002.5	69.9	1528	2 US-08-461-384B-6	Sequence 6, Appl1
19	4487.5	44.8	1527	2 US-09-647-140B-6	Sequence 6, Appl1
20	4487.5	44.8	1530	2 US-09-647-140B-33	Sequence 33, Appl1
21	3391.5	33.9	1503	2 US-09-647-140B-8	Sequence 8, Appl1
22	3385.5	33.8	1503	2 US-09-792-616-3	Sequence 3, Appl1
23	3218.5	32.1	1498	2 US-09-792-616-9	Sequence 9, Appl1
24	2374.5	23.7	1621	2 US-08-972-927-3	Sequence 3, Appl1
25	2349.5	23.5	1622	2 US-08-972-927-6	Sequence 6, Appl1
26	2266.5	22.6	1325	2 US-09-647-140B-2	Sequence 2, Appl1
27	2212.5	22.1	1464	2 US-10-012-896-1008	Sequence 1008, Ap

28	2195.5	21.9	1261	2 US-09-439-313-538	Sequence 538, App
29	2195.5	21.9	1261	2 US-09-636-215-538	Sequence 538, App
30	2195.5	21.9	1261	2 US-09-685-166A-538	Sequence 538, App
31	2195.5	21.9	1261	2 US-09-679-426-538	Sequence 538, App
32	2195.5	21.9	1261	2 US-09-759-143-538	Sequence 538, App
33	2195.5	21.9	1261	2 US-09-651-236-538	Sequence 538, App
34	2195.5	21.9	1261	2 US-09-657-279-538	Sequence 538, App
35	2195.5	21.9	1261	2 US-10-012-896-538	Sequence 538, App
36	2195.5	21.9	1261	2 US-10-012-896-1009	Sequence 1009, App
37	2118	21.1	1581	2 US-08-726-320-3	Sequence 3, Appl1
38	2118	21.1	1581	2 US-09-208-716-3	Sequence 3, Appl1
39	2115	21.1	1228	2 US-09-439-313-537	Sequence 537, App
40	2115	21.1	1228	2 US-09-636-215-537	Sequence 537, App
41	2115	21.1	1228	2 US-09-685-166A-537	Sequence 537, App
42	2115	21.1	1228	2 US-09-679-426-537	Sequence 537, App
43	2115	21.1	1228	2 US-09-759-143-537	Sequence 537, App
44	2115	21.1	1228	2 US-09-651-236-537	Sequence 537, App
45	2115	21.1	1228	2 US-09-657-279-537	Sequence 537, App

ALIGNMENTS

RESULT 1
US-08-463-092B-4
Sequence 4, Application US/08463092B
Patent No. 5766880
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: PARTO RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,092B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Streeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: Q1546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1531 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-463-092B-4

Query Match 78.5%; Score 7860; DB 1; Length 1531;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MALRGFCSADGSDPLMDMNTWTNTSNPDFTCFQNTVLVWPCFYLAACFPFFYLYSRH 60
DB 1 MALRGFCSADGSDPLMDMNTWTNTSNPDFTCFQNTVLVWPCFYLAACFPFFYLYSRH 60
QY 61 DRGIIQMTPLNKTKTALGFILMIYCMADLPYSPWERSRGIFLAVFLVSPILLGITLLA 120
DB 61 DRGIIQMTPLNKTKTALGFILMIYCMADLPYSPWERSRGIFLAVFLVSPILLGITLLA 120
QY 121 TFLQLERRKGVQSSGIMLTFMLVALYCALILRSKIMTALKEBAOYDLPRDITFYVYS 180
DB 121 TFLQLERRKGVQSSGIMLTFMLVALYCALILRSKIMTALKEBAOYDLPRDITFYVYS 180
QY 181 LLILQLVLSGSDSPLEFSETIHDNPPCBSSASFLSHITTTWTITGLIVRGIRPLEGSD 240
DB 181 LLILQLVLSGSDSPLEFSETIHDNPPCBSSASFLSHITTTWTITGLIVRGIRPLEGSD 240
QY 241 LMSLNKEDTSQOVPVLYQNMKKCAKTRKQPVVVVSSKDPAPKSSSKVDANEVEAL 300
DB 241 LMSLNKEDTSQOVPVLYQNMKKCAKTRKQPVVVVSSKDPAPKSSSKVDANEVEAL 300
QY 301 IVKSPOKEMNPSLFKVLYKTPGYFLMSFFPKALHDLMMFSGPOLIKLLIKFVNDTKAPD 360
DB 301 IVKSPOKEMNPSLFKVLYKTPGYFLMSFFPKALHDLMMFSGPOLIKLLIKFVNDTKAPD 360
QY 361 WQGFYTVLVFVTKLQTLVHQYFHICFVSGMRKTAIVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYTVLVFVTKLQTLVHQYFHICFVSGMRKTAIVIGAVYRKALVITNSARKSTV 420
QY 421 GEIYNLMSVDQRFMDLATTYINMIWSAPLOYILALYLLMLNLSGSLAGVAVMLMPVNV 480
DB 421 GEIYNLMSVDQRFMDLATTYINMIWSAPLOYILALYLLMLNLSGSLAGVAVMLMPVNV 480
QY 481 AVAMAKTKTYOVAMHKSQNRIRKLNNEILNGIKVLLKYAMELAFKDYLAIROBELKYLK 540
DB 481 AVAMAKTKTYOVAMHKSQNRIRKLNNEILNGIKVLLKYAMELAFKDYLAIROBELKYLK 540
QY 541 KSAVLSAVGTFTWCTPPLVALCTPAVYVTTIDENNILDAQTAFFVSLALFNILRPPLNLP 600
DB 541 KSAVLSAVGTFTWCTPPLVALCTPAVYVTTIDENNILDAQTAFFVSLALFNILRPPLNLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIRRPVKDGGNSITVRNATFTWARSDDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIRRPVKDGGNSITVRNATFTWARSDDPT 660
QY 661 LNCITFSIPBEGALVAVVGQVCGKSSLSALLAEMDKYEGHAIKGSVAVYPOQAIOND 720
DB 661 LNCITFSIPBEGALVAVVGQVCGKSSLSALLAEMDKYEGHAIKGSVAVYPOQAIOND 720
QY 721 SLRENILFGCOLBEPYRSVIOACALLPDLBITLBSGRTEIGEKGAVNSGGQKORVSLAR 780
DB 721 SLRENILFGCOLBEPYRSVIOACALLPDLBITLBSGRTEIGEKGAVNSGGQKORVSLAR 780
QY 781 AVYNSNADLYLEDDPLSAVDAVHGKHIENNVIQPKGMILKMKRILIVTHSMVLPQVDVIV 840
DB 781 AVYNSNADLYLEDDPLSAVDAVHGKHIENNVIQPKGMILKMKRILIVTHSMVLPQVDVIV 840
QY 841 MSGGKISMSGSYOELLARDGAFABELRTYASTEQDABENGCVTVSGPGKEAKOMENG 900
DB 841 MSGGKISMSGSYOELLARDGAFABELRTYASTEQDABENGCVTVSGPGKEAKOMENG 900
QY 901 LVYDSAGKOLQROLSSSSSYSGDISRRHNSYAELOKAEAKKEETWKLMEADKAQTQVYL 960
DB 901 LVYDSAGKOLQROLSSSSSYSGDISRRHNSYAELOKAEAKKEETWKLMEADKAQTQVYL 960

```

```

DB 901 LVYDSAGKOLQROLSSSSSYSGDISRRHNSYAELOKAEAKKEETWKLMEADKAQTQVYL 960
QY 961 SVYWDYKAIGLFISFSLFPMCNHVSALASNYMLSLMTDDPIVNGTOHTTKVRLSVYG 1020
DB 961 SVYWDYKAIGLFISFSLFPMCNHVSALASNYMLSLMTDDPIVNGTOHTTKVRLSVYG 1020
QY 1021 ALGISQIAVFGYSMAVSIIGIILASRCLHVDLLHSILRSWSPFERTPSGNLVNRSKEL 1080
DB 1021 ALGISQIAVFGYSMAVSIIGIILASRCLHVDLLHSILRSWSPFERTPSGNLVNRSKEL 1080
QY 1081 DTVDSMIPVITKFMGSLFENVIGACTYITLATPIAIIIPPLGIYFVQRPVYASSROL 1140
DB 1081 DTVDSMIPVITKFMGSLFENVIGACTYITLATPIAIIIPPLGIYFVQRPVYASSROL 1140
QY 1141 KRLESVRSRPVYSHFNETLIGVSYIRAFEOERFIHOSDLKVDENQAYPSIVANRMLA 1200
DB 1141 KRLESVRSRPVYSHFNETLIGVSYIRAFEOERFIHOSDLKVDENQAYPSIVANRMLA 1200
QY 1201 VRLCVCNCIYLPALPAVISRHSLSAGLVGLSVSYLQVTTYLNLVMSSEMETNIVA 1260
DB 1201 VRLCVCNCIYLPALPAVISRHSLSAGLVGLSVSYLQVTTYLNLVMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMIOETAPSSWPQYGRVFRVYCLRYEDLDVLRHINVTINGG 1320
DB 1261 VERLKEYSETEKAPMIOETAPSSWPQYGRVFRVYCLRYEDLDVLRHINVTINGG 1320
QY 1321 EKVGIVGRTAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTTIIPQDPVL 1380
DB 1321 EKVGIVGRTAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTTIIPQDPVL 1380
QY 1381 SGIIRMLNDPPSOYSDDEWNTSLAHLKDFVSALPKLDHBCAEGGENTSVGROLVCL 1440
DB 1381 SGIIRMLNDPPSOYSDDEWNTSLAHLKDFVSALPKLDHBCAEGGENTSVGROLVCL 1440
QY 1441 ARALRRTKTLVDEAFAVDELTDLLIOSTIRQFEDCTVLTARHLNTIMOTRYIVL 1500
DB 1441 ARALRRTKTLVDEAFAVDELTDLLIOSTIRQFEDCTVLTARHLNTIMOTRYIVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSAKADAGLV 1531
DB 1501 DKGEIOEYGAPSDLLQORGLFYSAKADAGLV 1531

```

RESULT 2
 US-08-462-109A-4
 Sequence 4, Application US/08462109A
 Patent No. 5882875
 GENERAL INFORMATION:
 APPLICANT: Cole, Susan P.C.
 APPLICANT: Deeley, Roger G.
 TITLE OF INVENTION: METHODS FOR IDENTIFYING
 TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHIVE & COCKFIELD
 STREET: 60 State Street, suite 510
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/462,109A
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/966,923
 FILING DATE: 27-OCT-1992
 APPLICATION NUMBER: 08/029,340

FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: P01-002CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-4

Query Match 78.5%; Score 7860; DB 1; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCADGSDPLMDMNTMTSNPDFTKCFONTVLVWVPCFYLAACPEFYFLYSRH 60
DB 1 MALRGFCADGSDPLMDMNTMTSNPDFTKCFONTVLVWVPCFYLAACPEFYFLYSRH 60
QY 61 DRGVIQMTPLNKTALGFLIMIVCMADLFYSFWERSGIFLAFVPLVSPTLGITLLA 120
DB 61 DRGVIQMTPLNKTALGFLIMIVCMADLFYSFWERSGIFLAFVPLVSPTLGITLLA 120
QY 121 TPLIQERRRGVSSGIMLTFMVALVCAALILSKIMTAKEDAOVDLPDITFYYPFS 180
DB 121 TPLIQERRRGVSSGIMLTFMVALVCAALILSKIMTAKEDAOVDLPDITFYYPFS 180
QY 121 TPLIQERRRGVSSGIMLTFMVALVCAALILSKIMTAKEDAOVDLPDITFYYPFS 180
DB 121 TPLIQERRRGVSSGIMLTFMVALVCAALILSKIMTAKEDAOVDLPDITFYYPFS 180
QY 181 LLLIQLVLSGFCSDSPLESETIHDNPNCPRESSASFLSRITFWMTGLIVGYRPLGSD 240
DB 181 LLLIQLVLSGFCSDSPLESETIHDNPNCPRESSASFLSRITFWMTGLIVGYRPLGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVVYSKDPAPKESKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVVYSKDPAPKESKVDANEVEAL 300
QY 301 IVSPQKEMNPSPLEKVLKTFGPFYFMSFFPKAIDHLMFSGPQILKLLKFNVDITAPD 360
DB 301 IVSPQKEMNPSPLEKVLKTFGPFYFMSFFPKAIDHLMFSGPQILKLLKFNVDITAPD 360
QY 361 WGGFYTVLLEFVACLOTVLVHOYFHLCPVSGMKIKTAVIGAVYRKALVITNSARKSSTV 420
DB 361 WGGFYTVLLEFVACLOTVLVHOYFHLCPVSGMKIKTAVIGAVYRKALVITNSARKSSTV 420
QY 421 GEIYVNLMSVDAQRFMDLATYINMISAPLQVITLALYLLMLNLPSPVLAVVAVMVPVN 480
DB 421 GEIYVNLMSVDAQRFMDLATYINMISAPLQVITLALYLLMLNLPSPVLAVVAVMVPVN 480
QY 481 AYVAMAKTKTYOVAMKSKDNRIKLMNITLNGIKYTKLYANELAFKDYALIROBELVLK 540
DB 481 AYVAMAKTKTYOVAMKSKDNRIKLMNITLNGIKYTKLYANELAFKDYALIROBELVLK 540
QY 541 KSAVLSAVGTFTWCTPEFLVALCTFAYVYVITDENNIIIDAOAFVSLAFNIIIRPELILP 600
DB 541 KSAVLSAVGTFTWCTPEFLVALCTFAYVYVITDENNIIIDAOAFVSLAFNIIIRPELILP 600
QY 601 MVISSIVQASVSLKRLRIFLSHELEBDSIEREPVKDGGTNSITVNAFTFVARSDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHELEBDSIEREPVKDGGTNSITVNAFTFVARSDPT 660
QY 661 LNTGTSIPRGALVAVVGVCGGSSLSLALAMDVEGHHVAIKGSVAVYVPOAMQND 720
DB 661 LNTGTSIPRGALVAVVGVCGGSSLSLALAMDVEGHHVAIKGSVAVYVPOAMQND 720
QY 721 SLRENILFGOLEEPYRSVYQACALLPDEILIPSGRTEIGEGVNLSSGQKQVSLAR 780

DB 721 SLRENILFGOLEEPYRSVYQACALLPDEILIPSGRTEIGEGVNLSSGQKQVSLAR 780
QY 781 AVYSNADITYFPDDPLSADVDAHVGKHIFENYIGPKMKLAKTRILLYTHSMSTPOVDYIV 840
DB 781 AVYSNADITYFPDDPLSADVDAHVGKHIFENYIGPKMKLAKTRILLYTHSMSTPOVDYIV 840
QY 841 MSGKISEMSYQELLRDQAFAPFLRTVASTQEOAEENGVTGVSQPGKAEAKOMNGM 900
DB 841 MSGKISEMSYQELLRDQAFAPFLRTVASTQEOAEENGVTGVSQPGKAEAKOMNGM 900
QY 901 LVYDSACKQOLQROSLSSSSYSYSDISRHNSTABLOKAEKKEETKLEADKQOTGVKL 960
DB 901 LVYDSACKQOLQROSLSSSSYSYSDISRHNSTABLOKAEKKEETKLEADKQOTGVKL 960
QY 961 SVYWDYKAIQGLFISFISIFLFCMNHVSALASNYWLSMTDDPIVNGTOHHTKRLSVYG 1020
DB 961 SVYWDYKAIQGLFISFISIFLFCMNHVSALASNYWLSMTDDPIVNGTOHHTKRLSVYG 1020
QY 1021 ALGISQIAVGYSMANVSTGIIASRCLHVDLHSLRSFMSFPERPSGNLVNREKEL 1080
DB 1021 ALGISQIAVGYSMANVSTGIIASRCLHVDLHSLRSFMSFPERPSGNLVNREKEL 1080
QY 1081 DTVDSMIPVYIKMFMSLFAVIGACIYIILATPIAIIIPPLGIYFFVORFYAASSROL 1140
DB 1081 DTVDSMIPVYIKMFMSLFAVIGACIYIILATPIAIIIPPLGIYFFVORFYAASSROL 1140
QY 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENQKAYVPSIVANWLA 1200
DB 1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENQKAYVPSIVANWLA 1200
QY 1201 VRLCVCNCIVLFAALPAVYSRHSLSAGLVGSVSVSLQVTTIYANLVMSSEMETNIVA 1260
DB 1201 VRLCVCNCIVLFAALPAVYSRHSLSAGLVGSVSVSLQVTTIYANLVMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVAFRNYCCRYREDDLPVLRHINVTING 1320
DB 1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVAFRNYCCRYREDDLPVLRHINVTING 1320
QY 1321 EKVIGVGRGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
DB 1321 EKVIGVGRGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
QY 1381 SGLRMLDLPFSQYSDSEVWTSLELAHKDPFVSLPKLDHECAEGGENSVGQROLVCL 1440
DB 1381 SGLRMLDLPFSQYSDSEVWTSLELAHKDPFVSLPKLDHECAEGGENSVGQROLVCL 1440
QY 1441 ARALLRKTILVDEATAVDLETDDLIOSTIRTOFEDCTVLTAHRLNTIMDYTRYIVL 1500
DB 1441 ARALLRKTILVDEATAVDLETDDLIOSTIRTOFEDCTVLTAHRLNTIMDYTRYIVL 1500
QY 1501 DKGEIOEGAPSDLLQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEGAPSDLLQORGLFYSMAKDAGLV 1531

RESULT 3
US-08-460-907B-4
Sequence 4, Application US/08460907B
Patent No. 5891724
GENERAL INFORMATION:
APPLICANT: Dealey, Roger G.
APPLICANT: Cole, Susan P.C.
TITLE OF INVENTION: METHODS FOR CONFERRING MULTIDRUG
RESISTANCE ON A CELL
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: PARTRO RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,907B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01551
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-907B-4

Query March 78.5%; Score 7860; DB 1; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSDGSDPLMDNNTWNTSNPDFTCFQNTVLMVPCFYLMACFPFYELYSRH 60
Db 1 MALRGFCSDGSDPLMDNNTWNTSNPDFTCFQNTVLMVPCFYLMACFPFYELYSRH 60
QY 61 DRGIYQMTPLNKTATGFLIMVCMADLFYSFMRSRGIFLAVFLVSPFLGLITLLA 120
Db 61 DRGIYQMTPLNKTATGFLIMVCMADLFYSFMRSRGIFLAVFLVSPFLGLITLLA 120
QY 121 TELIOLERRKVOSSGIMLTFMVLVAGLALIRSKIMTALKEQAQVULFPDITFYFYS 180
Db 121 TELIOLERRKVOSSGIMLTFMVLVAGLALIRSKIMTALKEQAQVULFPDITFYFYS 180
QY 181 LLLIQLVLSGSDSPLFSETIHDNPPCESSASFLSRITFMWITGLIVRGROPLESSD 240
Db 181 LLLIQLVLSGSDSPLFSETIHDNPPCESSASFLSRITFMWITGLIVRGROPLESSD 240
QY 241 LMSINKEDTSQVVPVVLVKNMKKECAKTRKQPVKVVYSSKDPAPQKSSKYDANEVEAL 300
Db 241 LMSINKEDTSQVVPVVLVKNMKKECAKTRKQPVKVVYSSKDPAPQKSSKYDANEVEAL 300
QY 301 IVKSPQKEMNPSLPRVLYKTGPRFLMSPFKAHIDLMFSGPOLIKLIRFVNDTKAPD 360
Db 301 IVKSPQKEMNPSLPRVLYKTGPRFLMSPFKAHIDLMFSGPOLIKLIRFVNDTKAPD 360
QY 361 WQGFYTYVLVFTVACTQTLVLAHOYFHIQFVSGMRIKTAVIGAVRKALVITNSARKSSTV 420
Db 361 WQGFYTYVLVFTVACTQTLVLAHOYFHIQFVSGMRIKTAVIGAVRKALVITNSARKSSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINIMISAPLQVILALYLLMLNLGSPVLAVAVMVLMEVN 480

Db 421 GEIVNLSVDAQRFMDLATYINIMISAPLQVILALYLLMLNLGSPVLAVAVMVLMEVN 480
QY 481 AVNMAKTKTYQVAMHMSKDRIRIKLMEIILNGIKVLKLYAMELAFKQVYLIRQBELVYLK 540
Db 481 AVNMAKTKTYQVAMHMSKDRIRIKLMEIILNGIKVLKLYAMELAFKQVYLIRQBELVYLK 540
QY 541 KSAVLSAVGFTVWCCTPFLVALCTFAVYVTIDENNIIIDAQTAVSALFNILRPPLNLP 600
Db 541 KSAVLSAVGFTVWCCTPFLVALCTFAVYVTIDENNIIIDAQTAVSALFNILRPPLNLP 600
QY 601 MVISIVQASVSLKRLRILFISHELEPDSIERRPVKGCGTNSITVNNATFTWARSDDPT 660
Db 601 MVISIVQASVSLKRLRILFISHELEPDSIERRPVKGCGTNSITVNNATFTWARSDDPT 660
QY 661 LMGITFSIPGALVAVAVGQVCGKSSLLSALLAEMDYEGHVALKGSVAIVPOQAMQND 720
Db 661 LMGITFSIPGALVAVAVGQVCGKSSLLSALLAEMDYEGHVALKGSVAIVPOQAMQND 720
QY 721 SLRENILFGCQLEBPYRSVYIQAACALLPDLEILPSGDRTEIGEXVNLSCGQKQVSLAR 780
Db 721 SLRENILFGCQLEBPYRSVYIQAACALLPDLEILPSGDRTEIGEXVNLSCGQKQVSLAR 780
QY 781 AVYSNADIYLFDDPLSAVDHVGKHIPEVNIIGPKMKKNTRIILVTHSMSTYLPQVDVYIV 840
Db 781 AVYSNADIYLFDDPLSAVDHVGKHIPEVNIIGPKMKKNTRIILVTHSMSTYLPQVDVYIV 840
QY 841 MSGGKISEMSYQELRLRDAFAEFLRTYASTEBOQDAENGVTVSGPGKEAKOMENGM 900
Db 841 MSGGKISEMSYQELRLRDAFAEFLRTYASTEBOQDAENGVTVSGPGKEAKOMENGM 900
QY 901 LVYDSAGKQIQRQLSSSSSYSGDISRRHNSYABLOKAEKKEETWKLMEADKAQTGVYL 960
Db 901 LVYDSAGKQIQRQLSSSSSYSGDISRRHNSYABLOKAEKKEETWKLMEADKAQTGVYL 960
QY 961 SVYWDYKKAIGLPISELSIFLPMCNHVSALASNYWLSLMTDDPLVNGTOEHTKRLSVYG 1020
Db 961 SVYWDYKKAIGLPISELSIFLPMCNHVSALASNYWLSLMTDDPLVNGTOEHTKRLSVYG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGIILASRCIAVDLHLSILRSPSPERTPSGNLVNRFSEKL 1080
Db 1021 ALGISOGIAVFGYSMAVSIIGIILASRCIAVDLHLSILRSPSPERTPSGNLVNRFSEKL 1080
QY 1081 DTVDSMIPEVYIKMFGSLFNVTIGACIVILLATPIAIIIPPLGLIYFVQRFVYASSROL 1140
Db 1081 DTVDSMIPEVYIKMFGSLFNVTIGACIVILLATPIAIIIPPLGLIYFVQRFVYASSROL 1140
QY 1141 KRLSEVSRSPPYSHFNFTLLGVSVYIRAFEBQERFIHQSDLKVDENQKAYPSIVANWLA 1200
Db 1141 KRLSEVSRSPPYSHFNFTLLGVSVYIRAFEBQERFIHQSDLKVDENQKAYPSIVANWLA 1200
QY 1201 VRLCEVNCIYLFALFAVISRHSLSAGLVGLSVSYSLQVTTYVNMIVRMSSEMETIIVA 1260
Db 1201 VRLCEVNCIYLFALFAVISRHSLSAGLVGLSVSYSLQVTTYVNMIVRMSSEMETIIVA 1260
QY 1261 VERLEKSETEKEAPMOIOETAPSSWPQVGRVEFRNYCLRYREDLDFVLRIHIVTINGG 1320
Db 1261 VERLEKSETEKEAPMOIOETAPSSWPQVGRVEFRNYCLRYREDLDFVLRIHIVTINGG 1320
QY 1321 EKVGIVRGTGAKSSLTGLFRINESAEGRIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
Db 1321 EKVGIVRGTGAKSSLTGLFRINESAEGRIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGLSRMNLDPESQVSDDEWMTSLELAHLKQFVSLPKLHCEKAGEGEMISVGOQOLVCL 1440
Db 1381 SGLSRMNLDPESQVSDDEWMTSLELAHLKQFVSLPKLHCEKAGEGEMISVGOQOLVCL 1440
QY 1441 ABALRLKTKILVDEEATAVVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDTRYIVL 1500
Db 1441 ABALRLKTKILVDEEATAVVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDTRYIVL 1500
QY 1501 DKGEIOEYGA PSDILIQORGLFYSMAKDAGLV 1531

Db 1501 DKGEI0EYGA5B5DLLQ0RGLFY5MAKDA6LV 1531

RESULT 4
US-08-463-179A-4
Sequence 4, Application US/08463179A
Patent No. 6001563
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Deoley, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,179A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Deconti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: PQI-002CP8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-179A-4

Query Match 78.5%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARGFSSAGSDPLMDWNTWNTSNPDFTKCFONTLVVWPCFYLMACPEFFLYSRH 60
Db 1 MARGFSSAGSDPLMDWNTWNTSNPDFTKCFONTLVVWPCFYLMACPEFFLYSRH 60
QY 61 DRGVIOWTPLNKTALGFLIMTYQADLFYSFWERSRGFLAPVPELVSTLTIGITLLA 120
Db 61 DRGVIOWTPLNKTALGFLIMTYQADLFYSFWERSRGFLAPVPELVSTLTIGITLLA 120
QY 121 TFLIQLERRRGVSSGIMLTFWLVAVCALAIRSKIMTALKEBAQVDFRDIITFYVFS 180
Db 121 TFLIQLERRRGVSSGIMLTFWLVAVCALAIRSKIMTALKEBAQVDFRDIITFYVFS 180
QY 181 LLLIIQLVLSGFSRSPFLFSTTHDPNCPRESSASFLSRITFWMTGLIVGYRPLRGSD 240
Db 181 LLLIIQLVLSGFSRSPFLFSTTHDPNCPRESSASFLSRITFWMTGLIVGYRPLRGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNWKKECAKTRQOPVAVVSSKDPQPKSSKVDANEVEAL 300

Db 241 LMSLNKEDTSEQVVPVLVKNWKKECAKTRQOPVAVVSSKDPQPKSSKVDANEVEAL 300
QY 301 IVKSPQKRNPSLFLVLYKTFPGYFLMSFFFKAIHDLMPSPQIILKLFVNDTAPD 360
Db 301 IVKSPQKRNPSLFLVLYKTFPGYFLMSFFFKAIHDLMPSPQIILKLFVNDTAPD 360
QY 361 MGGFYVLLFVTRACIOTLVHOFNHCPSGMRIRKTAIVGAVRKALVITNSARKSSTV 420
Db 361 MGGFYVLLFVTRACIOTLVHOFNHCPSGMRIRKTAIVGAVRKALVITNSARKSSTV 420
QY 421 GEIVNIMSVDQRFMDIATYINMIWSAPLOVILALYLMNLGFSVLAGVAVMLAMPVN 480
Db 421 GEIVNIMSVDQRFMDIATYINMIWSAPLOVILALYLMNLGFSVLAGVAVMLAMPVN 480
QY 481 AVNARKTKTYQVAHMSKDNRIKLMEINLINGIVALKYAMELAFKDVLAIRBELKVLK 540
Db 481 AVNARKTKTYQVAHMSKDNRIKLMEINLINGIVALKYAMELAFKDVLAIRBELKVLK 540
QY 541 KSAVLSAVGTFTWCTPELVALCTPAVYVITDENNILDQAFAVSLAFNLRPLNLP 600
Db 541 KSAVLSAVGTFTWCTPELVALCTPAVYVITDENNILDQAFAVSLAFNLRPLNLP 600
QY 601 MVISSIVQASVSLKRLIFLSHLEBPSIERRPVKDGCGTNSITVNNATFTMARSDPT 660
Db 601 MVISSIVQASVSLKRLIFLSHLEBPSIERRPVKDGCGTNSITVNNATFTMARSDPT 660
QY 661 LMGITPSIPBGALVAVVGQVCGKSSLSALLAEMDKVEGHVAKGSVAVYPOQAWIOND 720
Db 661 LMGITPSIPBGALVAVVGQVCGKSSLSALLAEMDKVEGHVAKGSVAVYPOQAWIOND 720
QY 721 SLEENILFGQLEBPYRSYIOACALLPDEIIPSGDREITGEKVNLSCGQORVSLAR 780
Db 721 SLEENILFGQLEBPYRSYIOACALLPDEIIPSGDREITGEKVNLSCGQORVSLAR 780
QY 781 AAVSNADIVFDDPLSAVDADVGHKIPENYIGPKGMKXNTRILYHMSMYLPOVDVYIV 840
Db 781 AAVSNADIVFDDPLSAVDADVGHKIPENYIGPKGMKXNTRILYHMSMYLPOVDVYIV 840
QY 841 MSGGKISEMGSYQELLARDGAFELRTYASTEOBDAENGVGVSGPKKAKOMENGM 900
Db 841 MSGGKISEMGSYQELLARDGAFELRTYASTEOBDAENGVGVSGPKKAKOMENGM 900
QY 901 LYTDSAGKQOLQROLSSSSSSGDISSRHNSATLQRAEAKKETWKLMEADKQOTGVKL 960
Db 901 LYTDSAGKQOLQROLSSSSSSGDISSRHNSATLQRAEAKKETWKLMEADKQOTGVKL 960
QY 961 SVYWDYMKAIGLFISFISFLFMCNHYVSALASNYMLSLMTDDPIVNGTOEHTKVRLSVG 1020
Db 961 SVYWDYMKAIGLFISFISFLFMCNHYVSALASNYMLSLMTDDPIVNGTOEHTKVRLSVG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGILASRCLHVDLHSLIRSPMSFFERTPSGNLVNRSKEL 1080
Db 1021 ALGISOGIAVFGYSMAVSIIGILASRCLHVDLHSLIRSPMSFFERTPSGNLVNRSKEL 1080
QY 1081 DTVDSMIPVYIKMFMSLENVIGACIVILLATPIAIIIPGLGIYFVQRFYVASSROL 1140
Db 1081 DTVDSMIPVYIKMFMSLENVIGACIVILLATPIAIIIPGLGIYFVQRFYVASSROL 1140
QY 1141 KRLSVSRSPVSHFNETLIGSVIRAFEOBFIHQSDLKVDENKAYYPSIVANRWLA 1200
Db 1141 KRLSVSRSPVSHFNETLIGSVIRAFEOBFIHQSDLKVDENKAYYPSIVANRWLA 1200
QY 1201 VRLCVCNCIVLPAALFAVYSRSLAGLVSVSYSLQVTTYLVNMLVMSSEMETNIVA 1260
Db 1201 VRLCVCNCIVLPAALFAVYSRSLAGLVSVSYSLQVTTYLVNMLVMSSEMETNIVA 1260
QY 1261 VERLKEYSSEKAPQIOETAPSSWPQVGRVEFNKYCLARRREDLDFVLRHINVTINGG 1320
Db 1261 VERLKEYSSEKAPQIOETAPSSWPQVGRVEFNKYCLARRREDLDFVLRHINVTINGG 1320
QY 1321 EKVGIVRGTAGSSLTGLGFRINESABGEIITDGINAKIGLHDFPKITIIPODPVLE 1380

Db 1321 EKVGIVGRTGAGKSSLTITGLFRINESAGEIITIDGINIAKIGLHDLRFKTIITIPQDPLV 1380
Qy 1381 SSGSRMNDPFSQXSDEEVTSLSLAHKDFVSALPDKLDHECAGSGENLSVGOROLVCL 1440
Db 1381 SSGSRMNDPFSQXSDEEVTSLSLAHKDFVSALPDKLDHECAGSGENLSVGOROLVCL 1440
Qy 1441 ARALLRKTKIIVLDBATNAVDLETDDLIOSTIRTOFEDCTVLTIAHRIINTMDYTRIVL 1500
Db 1441 ARALLRKTKIIVLDBATNAVDLETDDLIOSTIRTOFEDCTVLTIAHRIINTMDYTRIVL 1500
Qy 1501 DKGEIOEGAPSDLLQORGLFYMAKDGIV 1531
Db 1501 DKGEIOEGAPSDLLQORGLFYMAKDGIV 1531

RESULT 5

US-08-461-384B-4
Sequence 4, Application US/08461384B

Patent No. 6025473

GENERAL INFORMATION:

APPLICANT: Cole, Susan P.C.

APPLICANT: Dealey, Roger G.

TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESS: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS

STREET: Queen's University at Kingston

CITY: Kingston

STATE: Ontario

COUNTRY: CANADA

ZIP: K7L 3N6

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/461,384B

FILING DATE: 05-JUN-95

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/966,923

FILING DATE: 27-OCT-1992

APPLICATION NUMBER: 08/029,340

FILING DATE: 8-MAR-1993

APPLICATION NUMBER: 08/141,893

FILING DATE: 26-OCT-1993

APPLICATION NUMBER: 08/407,207

FILING DATE: 20-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Steeg, Carol Miernicki

REGISTRATION NUMBER: 39,539

REFERENCE/DOCKET NUMBER: 01547

TELECOMMUNICATION INFORMATION:

TELEPHONE: (613) 545-2342

TELEFAX: (613) 545-6853

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1531 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-461-384B-4

Query Match 78.5%; Score 7860; DB 2; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRFGCSADSGDPLMDNMVNTNSPDTFKCFONTVLWVPCFYIACFPFYFLYSRH 60
Db 1 MALRFGCSADSGDPLMDNMVNTNSPDTFKCFONTVLWVPCFYIACFPFYFLYSRH 60
Qy 61 DRGIQTPLNKTALGFLMLTVCWADLFYSFWRSGRIGFLAVPLVSPFLTGITTLA 120
Db 61 DRGIQTPLNKTALGFLMLTVCWADLFYSFWRSGRIGFLAVPLVSPFLTGITTLA 120

Db 61 DRGIQTPLNKTALGFLMLTVCWADLFYSFWRSGRIGFLAVPLVSPFLTGITTLA 120
Qy 121 TFLIQLERRKGVSSGIMLTFFWLVAIVCALAIRSKIMTALKEDAQVDLPFRDITFYVYS 180
Db 121 TFLIQLERRKGVSSGIMLTFFWLVAIVCALAIRSKIMTALKEDAQVDLPFRDITFYVYS 180
Qy 181 LLLIQLVLSGFSRDSPLFSETIHDNPPCPSSSASFLSRITFFWITGILVNGYROPLEGSD 240
Db 181 LLLIQLVLSGFSRDSPLFSETIHDNPPCPSSSASFLSRITFFWITGILVNGYROPLEGSD 240
Qy 241 LMSLNKEDTSEOVVPLVKMKKCECAKTRKOPVAVVYSSSDPAQPKSSKVDAAEEVVAL 300
Db 241 LMSLNKEDTSEOVVPLVKMKKCECAKTRKOPVAVVYSSSDPAQPKSSKVDAAEEVVAL 300
Qy 301 IVKSPQKEMNPJLFKVLKTFGPFYFLMSFFFKAIHDMFMFSGPQILKLLIKFVNDTRAPD 360
Db 301 IVKSPQKEMNPJLFKVLKTFGPFYFLMSFFFKAIHDMFMFSGPQILKLLIKFVNDTRAPD 360
Qy 361 WQGYFYTVLLFTVACLOTVLHOFYHICVPSGMRKTAIVIGAVYRKALVITNSARKSTV 420
Db 361 WQGYFYTVLLFTVACLOTVLHOFYHICVPSGMRKTAIVIGAVYRKALVITNSARKSTV 420
Qy 421 GEIVNLMSVDAORFMDLATYINMTWSAPLOVITLALYLMLNGPSVLAGVAVWVLAQPVN 480
Db 421 GEIVNLMSVDAORFMDLATYINMTWSAPLOVITLALYLMLNGPSVLAGVAVWVLAQPVN 480
Qy 481 AVAMAKTKTYOVAMKSKDNRIKLMNEILINGIKVLKLYAMELAFKDKVLAIRQELKVLK 540
Db 481 AVAMAKTKTYOVAMKSKDNRIKLMNEILINGIKVLKLYAMELAFKDKVLAIRQELKVLK 540
Qy 541 KSAVLSAVGFTFWCTPPLVALCTFPAVYVITDENNIIDAOCTAPVSLAFNLRPLNLP 600
Db 541 KSAVLSAVGFTFWCTPPLVALCTFPAVYVITDENNIIDAOCTAPVSLAFNLRPLNLP 600
Qy 601 MVISIVQASVSLKRLIFLSHERLEPDSIERRVKDGGSNTITVNAFTTARSDPT 660
Db 601 MVISIVQASVSLKRLIFLSHERLEPDSIERRVKDGGSNTITVNAFTTARSDPT 660
Qy 661 LMGITFPIPEGALVAVVGVQCGKSSLLSALLAEMDVGEHVALKGSVAVYVPOQAWIOND 720
Db 661 LMGITFPIPEGALVAVVGVQCGKSSLLSALLAEMDVGEHVALKGSVAVYVPOQAWIOND 720
Qy 721 SLRENILFGCOLLEPYRVSIVQACALLPDLILPDSGRTEIGEKVNLSSGOKORVSLAR 780
Db 721 SLRENILFGCOLLEPYRVSIVQACALLPDLILPDSGRTEIGEKVNLSSGOKORVSLAR 780
Qy 781 AVYSNADIVLEPDDLAVDAVGHIFENYVGPQMLKNKTRILVTHSMSTLPQVDVYIV 840
Db 781 AVYSNADIVLEPDDLAVDAVGHIFENYVGPQMLKNKTRILVTHSMSTLPQVDVYIV 840
Qy 841 MSGKISBMGSYQELIARDGAFAEFLRTYASTEQODAEENGVTGVSQPKAKOMENGM 900
Db 841 MSGKISBMGSYQELIARDGAFAEFLRTYASTEQODAEENGVTGVSQPKAKOMENGM 900
Qy 901 LVYDSAGKOLQROLSSSSSYSGDISRRHNSYAELOKAEAKKEFTWKLMEADKAOTGOVKL 960
Db 901 LVYDSAGKOLQROLSSSSSYSGDISRRHNSYAELOKAEAKKEFTWKLMEADKAOTGOVKL 960
Qy 961 SVYWDYMKATIGLFTSPLSIFLFCMNHVSALASNYWLSLMTDDPIVNGOETHTRLSYVG 1020
Db 961 SVYWDYMKATIGLFTSPLSIFLFCMNHVSALASNYWLSLMTDDPIVNGOETHTRLSYVG 1020
Qy 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLHLSILRSPMSFPERTPSGNLVNRFSKEL 1080
Db 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLHLSILRSPMSFPERTPSGNLVNRFSKEL 1080
Qy 1081 DTVDMSIPEVIMKMGSLFNVIGACIVILLATPIAIIIPPLGLIYFFVQGFVYASSROL 1140
Db 1081 DTVDMSIPEVIMKMGSLFNVIGACIVILLATPIAIIIPPLGLIYFFVQGFVYASSROL 1140
Qy 1141 KRLESVSRSPYSHFNETLGLGVSYIRAFEEQERFIHOSDLKVDENOKAYYSIVANRWLA 1200
Db 1141 KRLESVSRSPYSHFNETLGLGVSYIRAFEEQERFIHOSDLKVDENOKAYYSIVANRWLA 1200

QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVYSISLOVTTYLNMVLRMSSEMETNIVA 1260
DB 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVYSISLOVTTYLNMVLRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVFNRYCLARYEDLDFVLRHINVTINGG 1320
DB 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVFNRYCLARYEDLDFVLRHINVTINGG 1320
QY 1321 EKVGIVERTGAGKSLTLGLFRINESABGEIINDGINAKIGLHDLFKITIIIPQDPVLF 1380
DB 1321 EKVGIVERTGAGKSLTLGLFRINESABGEIINDGINAKIGLHDLFKITIIIPQDPVLF 1380
QY 1381 SGLSRNMLDPSQYSDSEBWTSLBLAKLKDPSALPKLDHECAGGENISVGOBOLVCL 1440
DB 1381 SGLSRNMLDPSQYSDSEBWTSLBLAKLKDPSALPKLDHECAGGENISVGOBOLVCL 1440
QY 1441 ARALLRKTILVLDEATAVADLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRIVL 1500
DB 1441 ARALLRKTILVLDEATAVADLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRIVL 1500
QY 1501 DKGEIQEYGA PSDLLQORGLFYSAKXAGLV 1531
DB 1501 DKGEIQEYGA PSDLLQORGLFYSAKXAGLV 1531

RESULT 6

US-09-647-140B-19
; Sequence 19, Application US/09647140B
; Patent No. 6803184
; GENERAL INFORMATION:
; APPLICANT: Fox Chase Cancer Center
; APPLICANT: Krush, Gary D.
; APPLICANT: Lee, Kun
; APPLICANT: Belinsky, Martin G.
; APPLICANT: Bain, Lisa J.
; TITLE OF INVENTION: MRP-Related ABC Transporter Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use Thereof
; FILE REFERENCE: PCCC 98-02
; CURRENT APPLICATION NUMBER: US/09/647,140B
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: PCT/US99/06644
; PRIOR FILING DATE: 1999-03-26
; PRIOR FILING DATE: 1998-03-27
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079,759
; PRIOR APPLICATION NUMBER: 60/095,153
; PRIOR FILING DATE: 1998-08-03
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-647-140B-19

Query Match 78.5%; Score 7860; DB 2; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSPILMDMNTYNTSNPDFTKCFONTLVVWPCFYLMACPFYFLYSR 60
DB 1 MALRGFCSADGSPILMDMNTYNTSNPDFTKCFONTLVVWPCFYLMACPFYFLYSR 60
QY 61 DRGYIOMTPINKTKTALGFLMTIVCMADLFYSPWERSRGIPLAPVFLVSPTLGITTLLA 120
DB 61 DRGYIOMTPINKTKTALGFLMTIVCMADLFYSPWERSRGIPLAPVFLVSPTLGITTLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFMWLVAVCALAIRSKIMTALKEDAOVDLFRDITFYVYS 180
DB 121 TFLIQLERRKGVSSGIMLTFMWLVAVCALAIRSKIMTALKEDAOVDLFRDITFYVYS 180
QY 181 LLLIQLVLSGFSRSPPLFSETIHDPNCPSSASFLSRITFMWITGLIVRGYQPLEGSD 240
DB 181 LLLIQLVLSGFSRSPPLFSETIHDPNCPSSASFLSRITFMWITGLIVRGYQPLEGSD 240

DB 181 LLLIQLVLSGFSRSPPLFSETIHDPNCPSSASFLSRITFMWITGLIVRGYQPLEGSD 240
QY 241 LMSLNKEDTSEOVVPUVYKWKKECAKTRQPKRVVYSSQDPAQPKSSKYDANEVEAL 300
DB 241 LMSLNKEDTSEOVVPUVYKWKKECAKTRQPKRVVYSSQDPAQPKSSKYDANEVEAL 300
QY 301 IVKSPQKKNPSLFKVLKTFEGPYFLMSPEFKAIHDMFMSGPOILKLLIKFVNDTAPD 360
DB 301 IVKSPQKKNPSLFKVLKTFEGPYFLMSPEFKAIHDMFMSGPOILKLLIKFVNDTAPD 360
QY 361 MCGYFTVLDFVTAQLOTVLHOYFHCIFVSGMRKIKTAVIGAVYRKALVITNSARKSTV 420
DB 361 MCGYFTVLDFVTAQLOTVLHOYFHCIFVSGMRKIKTAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLSVDAQRPMDLATTINNIWSPLOYIALLYMLNLGSPVLAGVAVMYLMPVN 480
DB 421 GEIVNLSVDAQRPMDLATTINNIWSPLOYIALLYMLNLGSPVLAGVAVMYLMPVN 480
QY 481 ANAMKTKTYQVAHMSKDNRIKLMEILNGIKVLKLYAMELAFKDKVLAIROBELKVLK 540
DB 481 ANAMKTKTYQVAHMSKDNRIKLMEILNGIKVLKLYAMELAFKDKVLAIROBELKVLK 540
QY 541 KSAVLSAVGTFTVCTPFLVALCTFAVYVITIDENNLDAQTAFVSIALFNILRPLNLP 600
DB 541 KSAVLSAVGTFTVCTPFLVALCTFAVYVITIDENNLDAQTAFVSIALFNILRPLNLP 600
QY 601 MVTSSIVQASVSLKRLRIFLSHEELPDSIERRPVKGGGTNSITRNATFTMARSPPT 660
DB 601 MVTSSIVQASVSLKRLRIFLSHEELPDSIERRPVKGGGTNSITRNATFTMARSPPT 660
QY 661 LMGITTSIPRGALVAVVGVCGCKSSLASLALMEMKVGEHVAIKGSVAVPOQAMQND 720
DB 661 LMGITTSIPRGALVAVVGVCGCKSSLASLALMEMKVGEHVAIKGSVAVPOQAMQND 720
QY 721 SLEENILFGQLEBEPYRSVIOACALLPDEILPSCGRTEIGKGVNLSGGQKORVSLAR 780
DB 721 SLEENILFGQLEBEPYRSVIOACALLPDEILPSCGRTEIGKGVNLSGGQKORVSLAR 780
QY 781 AVYSNADIVLPDPLSAVDHVGKHIPENYIGKGMKNKTRILVTHSMSYLPQVDYIV 840
DB 781 AVYSNADIVLPDPLSAVDHVGKHIPENYIGKGMKNKTRILVTHSMSYLPQVDYIV 840
QY 841 MSGGKISEMGSYVELLAPDAFAEPLRTYASTOEBDAENGVYGVSGPKKAKOMENGM 900
DB 841 MSGGKISEMGSYVELLAPDAFAEPLRTYASTOEBDAENGVYGVSGPKKAKOMENGM 900
QY 901 LVYDSAGKOLQRLSSSSSYSGDISRHNSTAELOKAEAKKETMKLMEADKQOTGVKL 960
DB 901 LVYDSAGKOLQRLSSSSSYSGDISRHNSTAELOKAEAKKETMKLMEADKQOTGVKL 960
QY 961 SVYWDYMKALGFLISFLSIFLFCNHNVSALASNYMLSLMTDDPIVNGTOEHTYVRLSVYG 1020
DB 961 SVYWDYMKALGFLISFLSIFLFCNHNVSALASNYMLSLMTDDPIVNGTOEHTYVRLSVYG 1020
QY 1021 ALGISGIAVFGSMVSVIGLILASRCLVVDLHSLIRSMSPFEETPSGNLVNRSKEL 1080
DB 1021 ALGISGIAVFGSMVSVIGLILASRCLVVDLHSLIRSMSPFEETPSGNLVNRSKEL 1080
QY 1081 DTVDSMIPEVIKMFMSLFFVIGACIVILLATPIAIIIPPLGLIYFFVORFYVASSROL 1140
DB 1081 DTVDSMIPEVIKMFMSLFFVIGACIVILLATPIAIIIPPLGLIYFFVORFYVASSROL 1140
QY 1141 KRLESVSRSPVYSHFNETLIGVSVIRAFEOERFIHQSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLESVSRSPVYSHFNETLIGVSVIRAFEOERFIHQSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVYSISLOVTTYLNMVLRMSSEMETNIVA 1260
DB 1201 VRLECVGNCIVLFAALFAVISRHSLSAGLVGLSVYSISLOVTTYLNMVLRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVFNRYCLARYEDLDFVLRHINVTINGG 1320
DB 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVFNRYCLARYEDLDFVLRHINVTINGG 1320

QY 1321 EKVGIVRTGAGKSLTGLFRINESABGEIIDGINIAKIGLHDLRPKTTIIPQDPVL 1380
| | | | |
Db 1321 EKVGIVRTGAGKSLTGLFRINESABGEIIDGINIAKIGLHDLRPKTTIIPQDPVL 1380
QY 1381 SGSLRMLDPPSOYSDDEWMTSLAHKDFVSALPKLDHECAGGENISVGOROLVCL 1440
| | | | |
Db 1381 SGSLRMLDPPSOYSDDEWMTSLAHKDFVSALPKLDHECAGGENISVGOROLVCL 1440
QY 1441 ARALRRTKILVDEATAVDELTDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRYVL 1500
| | | | |
Db 1441 ARALRRTKILVDEATAVDELTDDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRYVL 1500
QY 1501 DKGEIQEYGAESDILLQORGLFYSMADAGLV 1531
| | | | |
Db 1501 DKGEIQEYGAESDILLQORGLFYSMADAGLV 1531

RESULT 7
US-08-141-893-2
; Sequence 2, Application US/08141893
; Patent No. 5489519
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P. C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/141,893
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923; 08/029,340
; FILING DATE: 27-OCT-1992; 8-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A. Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: P01-002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5149
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-141-893-2

Query Match 78.4%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALNGFCADGSDPLMDNMVNTWNTSNDFPKCFONTVLVWPFCFYLAACFPFYFLYSRH 60
| | | | |
Db 1 MALNGFCADGSDPLMDNMVNTWNTSNDFPKCFONTVLVWPFCFYLAACFPFYFLYSRH 60
QY 61 DRGIOMPLNKTALGFLIMIVCMADLFYSFWRSGIFLAPVFLVSPLLGITTLLA 120
| | | | |
Db 61 DRGIOMPLNKTALGFLIMIVCMADLFYSFWRSGIFLAPVFLVSPLLGITTLLA 120

QY 121 TFLIOLERRKGVSSGIMLTFLVVALVCAIAIRSKIMTALKEDAOVDLFRDITFYFYS 180
| | | | |
Db 121 TFLIOLERRKGVSSGIMLTFLVVALVCAIAIRSKIMTALKEDAOVDLFRDITFYFYS 180
QY 181 LLLIQLVLSGFSRSPFLFSETIHDPNCPRESSASFSLRTFFMMITGLIVRGYQPLEGSD 240
| | | | |
Db 181 LLLIQLVLSGFSRSPFLFSETIHDPNCPRESSASFSLRTFFMMITGLIVRGYQPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNMKCECAKTRKQPVKVYVSSKDPAPKXSSRVDAANEVEAL 300
| | | | |
Db 241 LMSLNKEDTSEQVVPVLVKNMKCECAKTRKQPVKVYVSSKDPAPKXSSRVDAANEVEAL 300
QY 301 IVSPQKSNPSPFLKYLKTFPGYFLMSFFFKAIHDLMSGQQLKLLIKPVNDTAPD 360
| | | | |
Db 301 IVSPQKSNPSPFLKYLKTFPGYFLMSFFFKAIHDLMSGQQLKLLIKPVNDTAPD 360
QY 361 MOGFYTVLLFPTACQTLVLAHQYFHLCPVSGMRKTAVGAAYRKALVITNSARKSTV 420
| | | | |
Db 361 MOGFYTVLLFPTACQTLVLAHQYFHLCPVSGMRKTAVGAAYRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDAORFMDLATYINMIWSAPLQVILALYLLMLNLPVLAGVAVVLAHPVN 480
| | | | |
Db 421 GEIVNLMSVDAORFMDLATYINMIWSAPLQVILALYLLMLNLPVLAGVAVVLAHPVN 480
QY 481 AVMAKTKTYQVAHMSKDNRIKLANBIINGITVLYKLYAWELEFKDKVLAIRQELKVLK 540
| | | | |
Db 481 AVMAKTKTYQVAHMSKDNRIKLANBIINGITVLYKLYAWELEFKDKVLAIRQELKVLK 540
QY 541 KSAVLSAVGTFTWVCTPFLVALCTFAVYVITDENNIIDAOTAFVSLFNLIRPPLNLR 600
| | | | |
Db 541 KSAVLSAVGTFTWVCTPFLVALCTFAVYVITDENNIIDAOTAFVSLFNLIRPPLNLR 600
QY 601 MVISIIQVASVSLKRLRIFLSHELEBDSIERRPVKDGCGTNSITVANATFTWASDPPT 660
| | | | |
Db 601 MVISIIQVASVSLKRLRIFLSHELEBDSIERRPVKDGCGTNSITVANATFTWASDPPT 660
QY 661 LMGITFSIPGALVAVVGVQCGKSLSLALBMDKVBGHVAIKGSVAVYVPOQAMQND 720
| | | | |
Db 661 LMGITFSIPGALVAVVGVQCGKSLSLALBMDKVBGHVAIKGSVAVYVPOQAMQND 720
QY 721 SLRENILFQGLREPPYRSVIOACALLPDLLEIIPSGGRTEIGKGVNLSSGQKORVSLAR 780
| | | | |
Db 721 SLRENILFQGLREPPYRSVIOACALLPDLLEIIPSGGRTEIGKGVNLSSGQKORVSLAR 780
QY 781 AVYSNADIYLFDDPLSAVDAHVGKHIPEENVIGPKMKLNKTRILVTHSMSTYLPQVDYIIV 840
| | | | |
Db 781 AVYSNADIYLFDDPLSAVDAHVGKHIPEENVIGPKMKLNKTRILVTHSMSTYLPQVDYIIV 840
QY 841 MSGKISEMGSYOELLARDGAPAEFLTVASTBOEDAEENGVTGVSFGPKAKQEMNGM 900
| | | | |
Db 841 MSGKISEMGSYOELLARDGAPAEFLTVASTBOEDAEENGVTGVSFGPKAKQEMNGM 900
QY 901 LVYDSAKQIOLROUSSSSYSYSGDISRHNSYAEIQAKEAKKEETWKMEADKAQTGVKL 960
| | | | |
Db 901 LVYDSAKQIOLROUSSSSYSYSGDISRHNSYAEIQAKEAKKEETWKMEADKAQTGVKL 960
QY 961 SVYWDYMKAIQGLFISFISIFLFMCNHSALASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
| | | | |
Db 961 SVYWDYMKAIQGLFISFISIFLFMCNHSALASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISOGIAVFGYSMAVSGIILASRCLHVDLHLSILRSPMSFFERTPSGNLVNRFSEKL 1080
| | | | |
Db 1021 ALGISOGIAVFGYSMAVSGIILASRCLHVDLHLSILRSPMSFFERTPSGNLVNRFSEKL 1080
QY 1081 DTVDSMTPEVYIKMWSGLFNVIAGCIYIILATPIAIIIPPLGLIYFVQRFYASSROL 1140
| | | | |
Db 1081 DTVDSMTPEVYIKMWSGLFNVIAGCIYIILATPIAIIIPPLGLIYFVQRFYASSROL 1140
QY 1141 KRLSEVSRSPVYSHFNFTLLGVSVYIRAFEBEORERTHOSDLKVDENQAKYYSIYANRWLA 1200
| | | | |
Db 1141 KRLSEVSRSPVYSHFNFTLLGVSVYIRAFEBEORERTHOSDLKVDENQAKYYSIYANRWLA 1200
QY 1201 VRLCEVGNICIVLFAALFAVISRRHSLASGLVGLSVYSIQTVTYLNMLVRMSSEMETNIVA 1260
| | | | |

```

Db      1201  VRLCEVGNACIVLFAALPAVISRHSLSAGVLGVLSYSLSQVTTYNLVRMSSEMETIVA 1260
Qy      1261  VERKESESEKEKPMQIOETAPSSSWPQVRVFRNYCLARREDDLPVLRHINVTNGG 1320
Db      1261  VERKESESEKEKPMQIOETAPSSSWPQVRVFRNYCLARREDDLPVLRHINVTNGG 1320
Qy      1321  EKVGIVRTGAGKSLTLGLPRINESAGEIIGININAKIGLHDLRFKTIIPQDDVLF 1380
Db      1321  EKVGIVRTGAGKSLTLGLFRINESAGEIIGININAKIGLHDLRFKTIIPQDDVLF 1380
Qy      1381  SGLIRNMLDPFSQYSDSEVWTSLELAHKDFVSALPKLDHECAGEENLSVGQRLVCL 1440
Db      1381  SGLIRNMLDPFSQYSDSEVWTSLELAHKDFVSALPKLDHECAGEENLSVGQRLVCL 1440
Qy      1441  ARALLRRTKTLVDEATAVADLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRIVL 1500
Db      1441  ARALLRRTKTLVDEATAVADLETDLIQSTIRTOFEDCTVLTIAHRLNTIMDYTRIVL 1500
Qy      1501  DKGEIOEYGA PSDLQORGLFYSMAKDAGLV 1531
Db      1501  DKGEIOEYGA PSDLQORGLFYSMAKDAGLV 1531

```

```

RESULT 8
US-08-463-092B-2
; Sequence 2, Application US/08463092B
; Patent No. 576680
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,092B
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: Q1546
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853

```

```

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-092B-2

Query Match      78.4%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1  MALRGFCSADGSPPLMDMNTANTNSNDFKCFONTLVWVPCFYLMACFPFELYSRH 60
Db      1  MALRGFCSADGSPPLMDMNTANTNSNDFKCFONTLVWVPCFYLMACFPFELYSRH 60
Qy      61  DRGYIOWTPLNKTALGFLMWIVCAWDLFYFWMERSRGIFLAPVFLVSPTLGITTLA 120
Db      61  DRGYIOWTPLNKTALGFLMWIVCAWDLFYFWMERSRGIFLAPVFLVSPTLGITTLA 120
Qy      121  TPLIOLERRKGVSSGIMLTFWLVCAALAIRSKIMTALKEDAQVDLFRDITFYFYS 180
Db      121  TPLIOLERRKGVSSGIMLTFWLVCAALAIRSKIMTALKEDAQVDLFRDITFYFYS 180
Qy      181  LLLIOLVLSGFSRSPPLFSETIHDNPNCPSSASPLSRTIFWMTGLIVRGYQPLEGSD 240
Db      181  LLLIOLVLSGFSRSPPLFSETIHDNPNCPSSASPLSRTIFWMTGLIVRGYQPLEGSD 240
Qy      241  LMSLNKEDTSEQVVPVLVKNWKKCAKTRKQPKVYVSSKDPAPKSSKYDANEVEAL 300
Db      241  LMSLNKEDTSEQVVPVLVKNWKKCAKTRKQPKVYVSSKDPAPKSSKYDANEVEAL 300
Qy      301  IVKSPQKWNPSLPKVLVYKTPGYFLMSFFFAIHDLMFSGPQIILKLIKPNDTKAPD 360
Db      301  IVKSPQKWNPSLPKVLVYKTPGYFLMSFFFAIHDLMFSGPQIILKLIKPNDTKAPD 360
Qy      361  WQGYFTVLLFVYACQTLVHOYFHCFSGKRITAVI GAYYRALVITNSARSSIV 420
Db      361  WQGYFTVLLFVYACQTLVHOYFHCFSGKRITAVI GAYYRALVITNSARSSIV 420
Qy      421  GEIVNLMSYDAQRFMDLATYNNIMWAPLOVITLALYMLNLPVSLAGVAVWLMPVN 480
Db      421  GEIVNLMSYDAQRFMDLATYNNIMWAPLOVITLALYMLNLPVSLAGVAVWLMPVN 480
Qy      481  AVMAKTKTYQVAHMSKDNRIKLNEIINGIKVLKYAMELAFKDKVLAIRQEBELVKL 540
Db      481  AVMAKTKTYQVAHMSKDNRIKLNEIINGIKVLKYAMELAFKDKVLAIRQEBELVKL 540
Qy      541  KSAVYLSAVGFTFWVCPELVALCTFAVYVITDENNTILDQTAVALPFIILFPLNII 600
Db      541  KSAVYLSAVGFTFWVCPELVALCTFAVYVITDENNTILDQTAVALPFIILFPLNII 600
Qy      601  MVISSIVQASVSLKRIPLFSHELEPDSIERRPVQDGGTNSITTRNATFTARSDDPT 660
Db      601  MVISSIVQASVSLKRIPLFSHELEPDSIERRPVQDGGTNSITTRNATFTARSDDPT 660
Qy      661  LMGITFSPREGALVAVVGQVCGKSSLSALLAEMDKVEGHVAKISVAVVPOQAMIQND 720
Db      661  LMGITFSPREGALVAVVGQVCGKSSLSALLAEMDKVEGHVAKISVAVVPOQAMIQND 720
Qy      721  SURENITLFGQLEPFRYSVIOCALLPDEILIPSDRTEIGKGVNLSGGQKQVRVSLAR 780
Db      721  SURENITLFGQLEPFRYSVIOCALLPDEILIPSDRTEIGKGVNLSGGQKQVRVSLAR 780
Qy      781  AVYSNADITLFDPLSADVAHVKGKILFENVIGPKGLKXKTRILTVHSMSTYLPQVVI 840
Db      781  AVYSNADITLFDPLSADVAHVKGKILFENVIGPKGLKXKTRILTVHSMSTYLPQVVI 840
Qy      841  MSGGKISEMGSYQELLARDGAFAELRTYASTEOEDAEENGVTGSGGKEAKQWENG 900
Db      841  MSGGKISEMGSYQELLARDGAFAELRTYASTEOEDAEENGVTGSGGKEAKQWENG 900
Qy      901  LVTDAGKQLQRLSSSSYSQDISRHNHSTLQEAKEKKEBTWKLMBADKQTOGVTL 960

```

Db 901 LVYDSAKQLOROSSSSSSSGDSSRHNSRAEQKAEKEWTKMEADKAOTGVKL 960
Qy 961 SVYDYMKAIGLFTSFISIFLFCNCNVSALASNYWLSMTDPTVNGTQETKRLSYG 1020
Db 961 SVYDYMKAIGLFTSFISIFLFCNCNVSALASNYWLSMTDPTVNGTQETKRLSYG 1020
Qy 1021 ALGISOGIAVFGYMAVSIIGIILASRCIAVDLHLSILSPMSFFERTPSGNLVNFSKEL 1080
Db 1021 ALGISOGIAVFGYMAVSIIGIILASRCIAVDLHLSILSPMSFFERTPSGNLVNFSKEL 1080
Qy 1081 DTUNSMIPEVTKMWSLFNITGACIVILLATPIAIIIPPLGIYFFVQRFYVASSQOL 1140
Db 1081 DTUNSMIPEVTKMWSLFNITGACIVILLATPIAIIIPPLGIYFFVQRFYVASSQOL 1140
Qy 1141 KRLESVRSRPSVSHFNETLGVSVIRAFEBQERFHOSDLKVDENOKAYYSIVANRWLA 1200
Db 1141 KRLESVRSRPSVSHFNETLGVSVIRAFEBQERFHOSDLKVDENOKAYYSIVANRWLA 1200
Qy 1201 VRLBCVGNCLVLPALPAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVMSSEMETNIVA 1260
Db 1201 VRLBCVGNCLVLPALPAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVMSSEMETNIVA 1260
Qy 1261 VERLKEVSETEKAPWQIQETAPPSMPOVGRVEFRNCLRYREDLDFVLHINVTINGG 1320
Db 1261 VERLKEVSETEKAPWQIQETAPPSMPOVGRVEFRNCLRYREDLDFVLHINVTINGG 1320
Qy 1321 EKVGIVRTGAGKSLTGLFRINSAEGEIIIDGINAKIGLHDLRKITIIIPQDPLF 1380
Db 1321 EKVGIVRTGAGKSLTGLFRINSAEGEIIIDGINAKIGLHDLRKITIIIPQDPLF 1380
Qy 1381 SGLSRMNLDPSPQYSDSEEWTSLELAHLKDEVSALPDKLDEHCAEGENLSVGQROLVCL 1440
Db 1381 SGLSRMNLDPSPQYSDSEEWTSLELAHLKDEVSALPDKLDEHCAEGENLSVGQROLVCL 1440
Qy 1441 ARALLRKTIIVDEBATAVVDLETDLIQSTIRTOFECECTVLTIAHRLNTIMDTRVIVL 1500
Db 1441 ARALLRKTIIVDEBATAVVDLETDLIQSTIRTOFECECTVLTIAHRLNTIMDTRVIVL 1500
Qy 1501 DKGEIOEXGAPSDLLQORGLFYMAKADGLV 1531
Db 1501 DKGEIOEXGAPSDLLQORGLFYMAKADGLV 1531

RESULT 9

US-08-462-109A-2
Sequence 2, Application US/08462109A
Patent No. 5862875
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
APPLICANT: Deeley, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING
TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,109A
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992

APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: POI-002CP4
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-2

Query Match 78.4%; Score 7849; DB 1; Length 1531;

Best Local Similarity 99.9%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;

Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MALRGFCSADGSDPLMDMNTWNTSNPDFTKCFONTLVLVWPCFYLMACEPFYLYLSRH 60
Db 1 MALRGFCSADGSDPLMDMNTWNTSNPDFTKCFONTLVLVWPCFYLMACEPFYLYLSRH 60
Qy 61 DRGYIOWTPPLNKTKTALGFLIMVYCADLTFYSPERSRGIFLAPVFLVSPILGITTLLA 120
Db 61 DRGYIOWTPPLNKTKTALGFLIMVYCADLTFYSPERSRGIFLAPVFLVSPILGITTLLA 120
Qy 121 TFLQLERRRGVSGSGIMLFFWLVALYCALAIIKSKIMTALKEBQVDLFDRIFFYYFS 180
Db 121 TFLQLERRRGVSGSGIMLFFWLVALYCALAIIKSKIMTALKEBQVDLFDRIFFYYFS 180
Qy 181 LLLIQVLVSCFSDRSPLESETIHDNPNCPSSASFLSRITFWMITGLIVRGYRQPLEGSD 240
Db 181 LLLIQVLVSCFSDRSPLESETIHDNPNCPSSASFLSRITFWMITGLIVRGYRQPLEGSD 240
Qy 241 LMSLNKEDTSEQVVPVYVKNMKKECACTRKQPVYVYSSSDPAOPKSSSKYDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVYVKNMKKECACTRKQPVYVYSSSDPAOPKSSSKYDANEVEAL 300
Qy 301 IVKSPQKEMNPSPLEKVLVYKTFGPFYFLMSFFPKAIIHDLMPFSGPILKLLIKFVNDTAPD 360
Db 301 IVKSPQKEMNPSPLEKVLVYKTFGPFYFLMSFFPKAIIHDLMPFSGPILKLLIKFVNDTAPD 360
Qy 361 WQGYFYTVLLFVTAQTLVLAHQYFHCIVSGMRKITAIVGAVVRKALVITNSARKSSTV 420
Db 361 WQGYFYTVLLFVTAQTLVLAHQYFHCIVSGMRKITAIVGAVVRKALVITNSARKSSTV 420
Qy 421 GEIVNLSVDAQRPMIDLATYINMISAPLOYITALLYLMLNLGSPVLAGAVALMLAMPVN 480
Db 421 GEIVNLSVDAQRPMIDLATYINMISAPLOYITALLYLMLNLGSPVLAGAVALMLAMPVN 480
Qy 481 AVMAMKTKTQVAMKSKDNRIKLMNEILINGIKVLYKYAMELAKDKVLAIRBELKVLK 540
Db 481 AVMAMKTKTQVAMKSKDNRIKLMNEILINGIKVLYKYAMELAKDKVLAIRBELKVLK 540
Qy 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNNIIDAQTAFFSLALFNIRPLENLIP 600
Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNNIIDAQTAFFSLALFNIRPLENLIP 600
Qy 601 MVISSIVQSVSLKRLRIFLSHELEPDSIERRPVKGCGGTSITVNNATFTARSDEPT 660
Db 601 MVISSIVQSVSLKRLRIFLSHELEPDSIERRPVKGCGGTSITVNNATFTARSDEPT 660
Qy 661 LMGITFSPBGAIVAVVGVCGGSSLSLALAMDVEGVHAIKGSVAVYPOQAWIOND 720
Db 661 LMGITFSPBGAIVAVVGVCGGSSLSLALAMDVEGVHAIKGSVAVYPOQAWIOND 720

```

QY 721 SLRENIIFGCOLBEPYRSVIOACALLPDLIELPSGRTIEGKVNLSGGOKORVSLAR 780
| | | | |
DB 721 SLRENIIFGCOLBEPYRSVIOACALLPDLIELPSGRTIEGKVNLSGGOKORVSLAR 780
QY 781 AVYSNADIYLPDDPLSAVDAAVGHKIFENVIGPKMLKNKTRILVTHSMSTLPVDVYIV 840
| | | | |
DB 781 AVYSNADIYLPDDPLSAVDAAVGHKIFENVIGPKMLKNKTRILVTHSMSTLPVDVYIV 840
QY 841 MSGGKISEMSYOBLLARDGAFAPFLRTYASTBOBDAEENGVCVSGPGKEAQMENG 900
| | | | |
DB 841 MSGGKISEMSYOBLLARDGAFAPFLRTYASTBOBDAEENGVCVSGPGKEAQMENG 900
QY 901 LVTSAGKOLOROLSSSSSYSGDISRHNSYAELOKAEKEEYMKMEADKATOGVYL 960
| | | | |
DB 901 LVTSAGKOLOROLSSSSSYSGDISRHNSYAELOKAEKEEYMKMEADKATOGVYL 960
QY 961 SVYWDYKAIGLFISPLSIFLMCNHVSALASNYMSLMTDDPIVNGTOHTKVRSLVYG 1020
| | | | |
DB 961 SVYWDYKAIGLFISPLSIFLMCNHVSALASNYMSLMTDDPIVNGTOHTKVRSLVYG 1020
QY 1021 ALGISOGIAVPGYMANVIGGIIASRCLHVDLHSLRSPMSFEERTPSGMLVNRFSKEL 1080
| | | | |
DB 1021 ALGISOGIAVPGYMANVIGGIIASRCLHVDLHSLRSPMSFEERTPSGMLVNRFSKEL 1080
QY 1081 DTVDMSIPEVYKMGSLFNVIAGACIYILATPIAIIIPPLGLIYFFVORFYASSROL 1140
| | | | |
DB 1081 DTVDMSIPEVYKMGSLFNVIAGACIYILATPIAIIIPPLGLIYFFVORFYASSROL 1140
QY 1141 KRLESVSRSPYSHFNETLLGVSVYIRAFEBQERFIHOSDLKVDENOKAYPSIVANPMLA 1200
| | | | |
DB 1141 KRLESVSRSPYSHFNETLLGVSVYIRAFEBQERFIHOSDLKVDENOKAYPSIVANPMLA 1200
QY 1201 VRLCVCNACIVLPALPAVYISRHSLAGVLGVSYSLOQTTYNMIVRMSSEMETNIVA 1260
| | | | |
DB 1201 VRLCVCNACIVLPALPAVYISRHSLAGVLGVSYSLOQTTYNMIVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMOIOETAPSSMPOVGRVFRNYCARYREDLPLVLRHINVTNGG 1320
| | | | |
DB 1261 VERLKEYSETEKEAPMOIOETAPSSMPOVGRVFRNYCARYREDLPLVLRHINVTNGG 1320
QY 1321 EKVGIVRTGAGKSLTLGLFRINESAGEIILIDGINIAKIGLHDLRFKIIIPQDPVLF 1380
| | | | |
DB 1321 EKVGIVRTGAGKSLTLGLFRINESAGEIILIDGINIAKIGLHDLRFKIIIPQDPVLF 1380
QY 1381 SGIISRMNLDPPSOYSDEEVTSLBLAKDFVSALPKLDHECAGEGENTSVGQROLVCL 1440
| | | | |
DB 1381 SGIISRMNLDPPSOYSDEEVTSLBLAKDFVSALPKLDHECAGEGENTSVGQROLVCL 1440
QY 1441 ABALLRKTILVDEATAVDLETFDILIOSTIRPOFEDCTVLTAAHRLNTIMOTRYIVL 1500
| | | | |
DB 1441 ABALLRKTILVDEATAVDLETFDILIOSTIRPOFEDCTVLTAAHRLNTIMOTRYIVL 1500
QY 1501 DKGEIOEGAPSDLLQORGLFYSMAXDAGLV 1531
| | | | |
DB 1501 DKGEIOEGAPSDLLQORGLFYSMAXDAGLV 1531

```

```

ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,907B
FILING DATE: 05-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Mternicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: Q1551
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-907B-2

Query Match 78.4%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTWNTSNPDFTKCFONTLVVWPCFYLMACEPFYFLYLSRH 60
| | | | |
DB 1 MALRGFCSADGSDPLMDMNTWNTSNPDFTKCFONTLVVWPCFYLMACEPFYFLYLSRH 60
QY 61 DRGTYQMTPLNKTALGFLIMIVCWADLFYSFWERSRGIFLAPVPLVSTLLGITTLLA 120
| | | | |
DB 61 DRGTYQMTPLNKTALGFLIMIVCWADLFYSFWERSRGIFLAPVPLVSTLLGITTLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFMLVALVCALAIIRSKIMTALKXDAQVDLFRDITFYVFS 180
| | | | |
DB 121 TFLIQLERRKGVSSGIMLTFMLVALVCALAIIRSKIMTALKXDAQVDLFRDITFYVFS 180
QY 181 LLLIQLVLSGFSRSPFLFSETIHDPNPCPESSASFLSRIETMMITGLIYVGRQPLEGSD 240
| | | | |
DB 181 LLLIQLVLSGFSRSPFLFSETIHDPNPCPESSASFLSRIETMMITGLIYVGRQPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLYKWKKECAKTRQPKVYVSSDPAQPKESSKYVDANEVEAL 300
| | | | |
DB 241 LMSLNKEDTSEQVVPVLYKWKKECAKTRQPKVYVSSDPAQPKESSKYVDANEVEAL 300
QY 301 IVKSPQKEMNPSEFKVLYKTFGYPFLMSPEFKAIHDLMPGSGQILKLIKFNNDTKAPD 360
| | | | |
DB 301 IVKSPQKEMNPSEFKVLYKTFGYPFLMSPEFKAIHDLMPGSGQILKLIKFNNDTKAPD 360
QY 361 WQGYFTVLLFVTAQIOTVLAHQYFHCYVSGNRKITAIVGAAYRRALVITNSARKSSTV 420
| | | | |
DB 361 WQGYFTVLLFVTAQIOTVLAHQYFHCYVSGNRKITAIVGAAYRRALVITNSARKSSTV 420

```

421 GEIVNLSVDAQRPMDLATYINMIWSAPLOVILALYLLMLNLSGSVLAGVAVMVLMPVN 480
421 GEIVNLSVDAQRPMDLATYINMIWSAPLOVILALYLLMLNLSGSVLAGVAVMVLMPVN 480
481 AVAMAKTKYOVAMKSKDNRIKIMNELNIGIKYLKIYAMELAKRDKYLAIROSELKYLK 540
481 AVAMAKTKYOVAMKSKDNRIKIMNELNIGIKYLKIYAMELAKRDKYLAIROSELKYLK 540
541 KSAVLSAVGFTWCTPFLVALCTPAVYVTTIDENNILDAQAFAPSLALFNILRFPNLTP 600
541 KSAVLSAVGFTWCTPFLVALCTPAVYVTTIDENNILDAQAFAPSLALFNILRFPNLTP 600
601 MVISSIVQASVSLRLRILFISHELEPDSIRRRPVKOGGINSITVRNATFTWARSDEPT 660
601 MVISSIVQASVSLRLRILFISHELEPDSIRRRPVKOGGINSITVRNATFTWARSDEPT 660
661 LMGTFSTIPGALVAVVGVCCGKSLSLALLAEMDKVGHVALKGSVAVYVPOQAMQND 720
661 LMGTFSTIPGALVAVVGVCCGKSLSLALLAEMDKVGHVALKGSVAVYVPOQAMQND 720
721 SLRENIIFGCOLLEPPYRSVIOACALLPDLILPSGRTIEGKGVNLSGGOKORVSLAR 780
721 SLRENIIFGCOLLEPPYRSVIOACALLPDLILPSGRTIEGKGVNLSGGOKORVSLAR 780
781 AVYSNADIYLFDDPLSAVDAAVGHKIFENVIQPKGMLKNKTRILIVTHSMSTYLPQVDVIL 840
781 AVYSNADIYLFDDPLSAVDAAVGHKIFENVIQPKGMLKNKTRILIVTHSMSTYLPQVDVIL 840
841 MSGGKISEMGSYOELLARDGAFAPFLRYASTEOBQDAENGVGVSGPEAKQEMEGM 900
841 MSGGKISEMGSYOELLARDGAFAPFLRYASTEOBQDAENGVGVSGPEAKQEMEGM 900
901 LVTSAGAKOLOROLSSSSSSSGDISRHHNSTAELOKAKKEETWKLMEADKAOQGVKL 960
901 LVTSAGAKOLOROLSSSSSSSGDISRHHNSTAELOKAKKEETWKLMEADKAOQGVKL 960
961 SVYWDYMKAIQLFISFLSIFLMFCNHSALASNYWLSMTDPIVNGTQEHTRVLSYVG 1020
961 SVYWDYMKAIQLFISFLSIFLMFCNHSALASNYWLSMTDPIVNGTQEHTRVLSYVG 1020
1021 ALGISOGIAVFGYMAVSIIGIILASRCHADVLHLSILSPMSFPERTSSGULVNRFSKL 1080
1021 ALGISOGIAVFGYMAVSIIGIILASRCHADVLHLSILSPMSFPERTSSGULVNRFSKL 1080
1081 DTVDSMIPEVIMKMGSLFNVIACIYVLLATPIAIIIPPLGIYFVQRPYVASSRQL 1140
1081 DTVDSMIPEVIMKMGSLFNVIACIYVLLATPIAIIIPPLGIYFVQRPYVASSRQL 1140
1141 KRLBSVSRSPYSHFNETLLGVSVIRAFEBQERFIHQSDLVKVDENOKAYPSIVANRWLA 1200
1141 KRLBSVSRSPYSHFNETLLGVSVIRAFEBQERFIHQSDLVKVDENOKAYPSIVANRWLA 1200
1201 VRLCEVGNICIVLPALPAFVIRSHSLAGLVGLSVSYSLQVTTYLNMVLRMSSEMETNIVA 1260
1201 VRLCEVGNICIVLPALPAFVIRSHSLAGLVGLSVSYSLQVTTYLNMVLRMSSEMETNIVA 1260
1261 VERLKEVSETEKEAPMOIQETRPSPSPQVGRVERRNCLEKREDDLVLRHINVTINGG 1320
1261 VERLKEVSETEKEAPMOIQETRPSPSPQVGRVERRNCLEKREDDLVLRHINVTINGG 1320
1321 EKVGIVGRTGAGKSLTJGLFRINSAGEIITDGINIAKIGLHDLRFKTIITIPDPVLF 1380
1321 EKVGIVGRTGAGKSLTJGLFRINSAGEIITDGINIAKIGLHDLRFKTIITIPDPVLF 1380
1381 SGSLRMLNDPSPQYSDEEVMVTSLELAHLKDFVSALPDYLDHCEAGGNNLSVGQOLVCL 1440
1381 SGSLRMLNDPSPQYSDEEVMVTSLELAHLKDFVSALPDYLDHCEAGGNNLSVGQOLVCL 1440
1441 ABALLRKTIIVLEBATAVLETTDDLOSTIRTOFEDCTVLTIAHRLNTIMDYRIVVL 1500
1441 ABALLRKTIIVLEBATAVLETTDDLOSTIRTOFEDCTVLTIAHRLNTIMDYRIVVL 1500
1501 DKGEIOEYGAPSDLLQORGLFYSMKADAGLV 1531
1501 DKGEIOEYGAPSDLLQORGLFYSMKADAGLV 1531

Db 1501 DKGEIOEYGAPSDLLQORGLFYSMKADAGLV 1531
RESULT 11
US-08-463-179A-2
Sequence 2, Application US/08463179A
Patent No. 6001563
GENERAL INFORMATION:
APPLICANT: Cole, Susan P.C.
TITLE OF INVENTION: METHODS FOR IDENTIFYING CHEMOSENSITIZERS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,179A
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Decont, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: P01-002CP8
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-5941
TELEFAX: (617) 227-7400
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-179A-2
Query Match 78.4%; Score 7849; DB 2; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MALGFCASDSDLLMNMVNTWNTSNPDFTKCFONTLVWVPCFYLMACPEFYLYLSRH 60
Db 1 MALGFCASDSDLLMNMVNTWNTSNPDFTKCFONTLVWVPCFYLMACPEFYLYLSRH 60
QY 61 DRGIQMTPLNKTKTALGFLIMVCMADLFYSFERSRGIFLAVFLVSPTLGITTLLA 120
Db 61 DRGIQMTPLNKTKTALGFLIMVCMADLFYSFERSRGIFLAVFLVSPTLGITTLLA 120
QY 121 TFLIQLEBRKGVQSSGIMLTFMWLVAVCALAIIRSKIMTALKEDAQVDLFFDITFYVFS 180
Db 121 TFLIQLEBRKGVQSSGIMLTFMWLVAVCALAIIRSKIMTALKEDAQVDLFFDITFYVFS 180
QY 121 TFLIQLEBRKGVQSSGIMLTFMWLVAVCALAIIRSKIMTALKEDAQVDLFFDITFYVFS 180
Db 121 TFLIQLEBRKGVQSSGIMLTFMWLVAVCALAIIRSKIMTALKEDAQVDLFFDITFYVFS 180
QY 181 LLLIQVLVSCFSDSPFLFSETIHDNPNCPRESSASFLRITTWMTTGLIVRGYRQPLEGSD 240
Db 181 LLLIQVLVSCFSDSPFLFSETIHDNPNCPRESSASFLRITTWMTTGLIVRGYRQPLEGSD 240
QY 181 LLLIQVLVSCFSDSPFLFSETIHDNPNCPRESSASFLRITTWMTTGLIVRGYRQPLEGSD 240
Db 181 LLLIQVLVSCFSDSPFLFSETIHDNPNCPRESSASFLRITTWMTTGLIVRGYRQPLEGSD 240

QY 241 LMSLNKEDTSEOVVPLVKNWKECAKTRKQPVKNVYSKDPQPKSSKYDANEVEAL 300
DB 241 LMSLNKEDTSEOVVPLVKNWKECAKTRKQPVKNVYSKDPQPKSSKYDANEVEAL 300
QY 301 IVKSPQKEMNPISLFKVLKTFGPRFLMSFFPKAIDHLMFSGPOILKLIKPVNDYAPD 360
DB 301 IVKSPQKEMNPISLFKVLKTFGPRFLMSFFPKAIDHLMFSGPOILKLIKPVNDYAPD 360
QY 361 MOGYFYVLLFVNACTOTVLHOYFHI CFSGMBIKTAVIGAVYRKALVITNSRKSTV 420
DB 361 MOGYFYVLLFVNACTOTVLHOYFHI CFSGMBIKTAVIGAVYRKALVITNSRKSTV 420
QY 421 GEIVNLMASVDAORFMDLATYINNMISAPLQVITLALYLLMNTGPSVLAGVAVMYLMPVN 480
DB 421 GEIVNLMASVDAORFMDLATYINNMISAPLQVITLALYLLMNTGPSVLAGVAVMYLMPVN 480
QY 481 AVAMAKTKTYQVAHMSKDNRIKLMNEILNGIKVLKYAMELAFKDVLAIRQBELKVLK 540
DB 481 AVAMAKTKTYQVAHMSKDNRIKLMNEILNGIKVLKYAMELAFKDVLAIRQBELKVLK 540
QY 541 KSAVLSAVGTFTWCTPFLVATCTPAVYVITDENNIIIDQTAFAVSLAFNIIARPLNIIIP 600
DB 541 KSAVLSAVGTFTWCTPFLVATCTPAVYVITDENNIIIDQTAFAVSLAFNIIARPLNIIIP 600
QY 601 MVISSIVQASVSLKRLIFLSHEELBDSIERRPVKDGGGNSITVNAFTMARSDPT 660
DB 601 MVISSIVQASVSLKRLIFLSHEELBDSIERRPVKDGGGNSITVNAFTMARSDPT 660
QY 661 LNCITPSIPBGAIVAVVGVGCCGSLLSALLAEMDVKGHVAIKGSVAVVPOQAWIOND 720
DB 661 LNCITPSIPBGAIVAVVGVGCCGSLLSALLAEMDVKGHVAIKGSVAVVPOQAWIOND 720
QY 721 SLRENIIFGQOLBEPYRYSYIOACALLPDLFILPSGRTIEGEGVNLSCGOKORVSLAR 780
DB 721 SLRENIIFGQOLBEPYRYSYIOACALLPDLFILPSGRTIEGEGVNLSCGOKORVSLAR 780
QY 781 AAVSNADIYLPDPLSAVDHVGKHIPEVYIGPKMKKNTRIILVTHSMGYLPQVDYIIV 840
DB 781 AAVSNADIYLPDPLSAVDHVGKHIPEVYIGPKMKKNTRIILVTHSMGYLPQVDYIIV 840
QY 841 MSGGKISEMGSYOELLARDGAFELRTYASTEOQDAEENGVYSGPGKEAKOMENGM 900
DB 841 MSGGKISEMGSYOELLARDGAFELRTYASTEOQDAEENGVYSGPGKEAKOMENGM 900
QY 901 LVNDSAGKOLQROLSSSSSYSGDISRRHNSTAEIOKAEKKEEYKLMEDAKQOTGVKL 960
DB 901 LVNDSAGKOLQROLSSSSSYSGDISRRHNSTAEIOKAEKKEEYKLMEDAKQOTGVKL 960
QY 961 SVYWDYKAIKGLIFSLIFLFCMCHVSALASNYWLSLMTDDPIVNGTOEHTKVRLSVYG 1020
DB 961 SVYWDYKAIKGLIFSLIFLFCMCHVSALASNYWLSLMTDDPIVNGTOEHTKVRLSVYG 1020
QY 1021 ALGISOGIAVFGYSMAVISGIIILASRCLAHVLDHLSIRSPMSFEERTPSGNLVNRFSEKL 1080
DB 1021 ALGISOGIAVFGYSMAVISGIIILASRCLAHVLDHLSIRSPMSFEERTPSGNLVNRFSEKL 1080
QY 1081 DTVDSMIPETIKPMFMSLFNVIGACTYIILATPIAIIIPPLGLIYFVORFYVASSROL 1140
DB 1081 DTVDSMIPETIKPMFMSLFNVIGACTYIILATPIAIIIPPLGLIYFVORFYVASSROL 1140
QY 1141 KRLSVSRSPVYSHFNFTLLGVSVIRAFEOERIHOSDLKVDENQAYPSIVANRWLA 1200
DB 1141 KRLSVSRSPVYSHFNFTLLGVSVIRAFEOERIHOSDLKVDENQAYPSIVANRWLA 1200
QY 1201 VRLVCVNCIIVLPAALFAVISRHSLSAGVLGSYSVLQVTTYLMLVMSSEMETNIVA 1260
DB 1201 VRLVCVNCIIVLPAALFAVISRHSLSAGVLGSYSVLQVTTYLMLVMSSEMETNIVA 1260
QY 1261 VERLKESETTEKAPMOIOETAPSSSPQVGRVFRNYCLRYREDLDFVLRHINVTINGG 1320
DB 1261 VERLKESETTEKAPMOIOETAPSSSPQVGRVFRNYCLRYREDLDFVLRHINVTINGG 1320
QY 1321 EKVGIVGRTAGKGSLLTGLFRINESAGEIIDIINIAKIGLHDLRFKTIIPQDPVL 1380

DB 1321 EKVGIVGRTAGKGSLLTGLFRINESAGEIIDIINIAKIGLHDLRFKTIIPQDPVL 1380
QY 1381 SCSLRNMLDPFSQYSBEWVTSLELAHKDVFSLPDKLDHECAEGGENISVQOROLVCL 1440
DB 1381 SCSLRNMLDPFSQYSBEWVTSLELAHKDVFSLPDKLDHECAEGGENISVQOROLVCL 1440
QY 1441 ARALLKRTKILVDEBATAVDETDLIQSTIRTOPEDCVLTIAHRLNTIMDYRTIYL 1500
DB 1441 ARALLKRTKILVDEBATAVDETDLIQSTIRTOPEDCVLTIAHRLNTIMDYRTIYL 1500
QY 1501 DKGEIOEYGA PSDILOQRGLFYSMAKDAGIV 1531
DB 1501 DKGEIOEYGA PSDILOQRGLFYSMAKDAGIV 1531

RESULT 12
US-08-461-384B-2
; Sequence 2, Application US/08461384B
; Patent No. 6025473
; GENERAL INFORMATION:
; APPLICANT: Cole, Susan P.C.
; APPLICANT: Deeley, Roger G.
; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PARTIQ RESEARCH & DEVELOPMENT INNOVATIONS
; STREET: Queen's University at Kingston
; CITY: Kingston
; STATE: Ontario
; COUNTRY: CANADA
; ZIP: K7L 3N6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,384B
; FILING DATE: 05-JUN-95
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/966,923
; FILING DATE: 27-OCT-1992
; APPLICATION NUMBER: 08/029,340
; FILING DATE: 8-MAR-1993
; APPLICATION NUMBER: 08/141,893
; FILING DATE: 26-OCT-1993
; APPLICATION NUMBER: 08/407,207
; FILING DATE: 20-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Steeg, Carol Miernicki
; REGISTRATION NUMBER: 39,539
; REFERENCE/DOCKET NUMBER: Q1547
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 545-2342
; TELEFAX: (613) 545-6853
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1531 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-384B-2

Query Match 78.4%; Score 7849; DB 2; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MALRGCSADSGSPILMDANVTWNTSNPDPKTCQNTVLVWVPCFYIMACFPFFLYLSRH 60
DB 1 MALRGCSADSGSPILMDANVTWNTSNPDPKTCQNTVLVWVPCFYIMACFPFFLYLSRH 60
QY 61 DRGYQMTPLNKTKTALGLFLMWVCWADLFYSPWERSRGIPLAPFLVPSVTLTGITTLA 120

Db 61 DRGIQMTPIKNTKTAAGFLMTIWMADLFYSFWEBSRGIFLAVFVLSPLLGIITLLA 120
Qy 121 TFLIQLERRKGVSSGIMLFEMLVALCALAILRSKIMTALKEDAOYDLFPHDITFYFYS 180
Db 121 TFLIQLERRKGVSSGIMLFTWVALCALAILRSKIMTALKEDAOYDLFPHDITFYFYS 180
Qy 181 LLLIQLVLCFSDRSPLFSETHIDPNPCPESSASFLSRITFEWITGLIVRGYROPLEGSD 240
Db 181 LLLIQLVLCFSDRSPLFSETHIDPNPCPESSASFLSRITFEWITGLIVRGYROPLEGSD 240
Qy 241 LMSLNKEDTSQOVVPVLYKMKKECAKTRKOPVYKVSXKDPAPKSSSKTDANEVBAL 300
Db 241 LMSLNKEDTSQOVVPVLYKMKKECAKTRKOPVYKVSXKDPAPKSSSKTDANEVBAL 300
Qy 301 IVKSPQKEMNPSLFKVLKTFEGFELMSFFPKALHDLMMFSGPOLIKLLIFVNDTAPD 360
Db 301 IVKSPQKEMNPSLFKVLKTFEGFELMSFFPKALHDLMMFSGPOLIKLLIFVNDTAPD 360
Qy 361 WQGYFTVLLFVTAQCLQTLVLHGYFHI CFVSGMR IKTAVIGAVYRKALVITNSARKSTV 420
Db 361 WQGYFTVLLFVTAQCLQTLVLHGYFHI CFVSGMR IKTAVIGAVYRKALVITNSARKSTV 420
Qy 421 GEIYNLMSVDAQRMDLATTYINMTWSAPLOYITLALYLMNLGSPVLAGVAVMLAMPVN 480
Db 421 GEIYNLMSVDAQRMDLATTYINMTWSAPLOYITLALYLMNLGSPVLAGVAVMLAMPVN 480
Qy 481 AVNMAKTKTYOVAMKSKDNRIKLMNEILNGIKYKLYAMELAPDKVLAIROBELKYLK 540
Db 481 AVNMAKTKTYOVAMKSKDNRIKLMNEILNGIKYKLYAMELAPDKVLAIROBELKYLK 540
Qy 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNIIDAQTAFLALFNILREPLNLP 600
Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNIIDAQTAFLALFNILREPLNLP 600
Qy 601 MVISSIYQASVSLKRLRIFLSHEBELPDSIRRPVKOGGNGNSTVNNATTTWASDPPT 660
Db 601 MVISSIYQASVSLKRLRIFLSHEBELPDSIRRPVKOGGNGNSTVNNATTTWASDPPT 660
Qy 661 LINGITFSIPGALVAVVQVCGCKSSLLSALLAEMDKVEGHVALKGSAAVYPOQAMIOND 720
Db 661 LINGITFSIPGALVAVVQVCGCKSSLLSALLAEMDKVEGHVALKGSAAVYPOQAMIOND 720
Qy 721 SLRENILFGCQLEBPYRSVIOACALLPDLILPSGRTIEGKGVNLSSGQOKORVSLAR 780
Db 721 SLRENILFGCQLEBPYRSVIOACALLPDLILPSGRTIEGKGVNLSSGQOKORVSLAR 780
Qy 781 AVYNNADIIYLFDDPLSAVDHVGKHIFFENVIGPKGMLKNKTRILVTHSMSTLPOVDVYIV 840
Db 781 AVYNNADIIYLFDDPLSAVDHVGKHIFFENVIGPKGMLKNKTRILVTHSMSTLPOVDVYIV 840
Qy 841 MSGGKISEMGSYOEILLARDGAFAEFLRTYASTBOEODAEENGAVGVSFGKEAKOMENGM 900
Db 841 MSGGKISEMGSYOEILLARDGAFAEFLRTYASTBOEODAEENGAVGVSFGKEAKOMENGM 900
Qy 901 LVNTDSAGKOLORQLSSSSSYSGDISRHNSSTAELQKAEKKEETWKLMEADKAOTGOYKL 960
Db 901 LVNTDSAGKOLORQLSSSSSYSGDISRHNSSTAELQKAEKKEETWKLMEADKAOTGOYKL 960
Qy 961 SVYWDYMKAIQGLFISFLSIFLPMCNHVSALASNTWLSLMTDDPIYNGTQHTKYLRSYVG 1020
Db 961 SVYWDYMKAIQGLFISFLSIFLPMCNHVSALASNTWLSLMTDDPIYNGTQHTKYLRSYVG 1020
Qy 1021 ALGISOGIAVGVMAVSIIGIILASRCIADVLHSIILSPMSFPERTSGLVNRFSKEL 1080
Db 1021 ALGISOGIAVGVMAVSIIGIILASRCIADVLHSIILSPMSFPERTSGLVNRFSKEL 1080
Qy 1081 DTVDMSIPEVIKMGSLFNIVIGACTIVLLATPIAIIIPGLIYFFVQGFYVASSROL 1140
Db 1081 DTVDMSIPEVIKMGSLFNIVIGACTIVLLATPIAIIIPGLIYFFVQGFYVASSROL 1140
Qy 1141 KRLSVSRSPVYSHPNETLLGVSVIRAFEBQERFIHOSDLKVDENQKAYPSIVANRWLA 1200
Db 1141 KRLSVSRSPVYSHPNETLLGVSVIRAFEBQERFIHOSDLKVDENQKAYPSIVANRWLA 1200

Db 1141 KRLSVSRSPVYSHPNETLLGVSVIRAFEBQERFIHOSDLKVDENQKAYPSIVANRWLA 1200
Qy 1201 VRLCEVGNCTIVLPAALFAVVISRHSLSAGLVGLSVSYSLQVTTYLNMVLRMSSEMETNIVA 1260
Db 1201 VRLCEVGNCTIVLPAALFAVVISRHSLSAGLVGLSVSYSLQVTTYLNMVLRMSSEMETNIVA 1260
Qy 1261 VERLKEYSETEKAPMWOIOETAPSSWPQYGRVBFRRNYCIRYEDDLFVLRHINVTINGG 1320
Db 1261 VERLKEYSETEKAPMWOIOETAPSSWPQYGRVBFRRNYCIRYEDDLFVLRHINVTINGG 1320
Qy 1321 EKVGIVRTAGKSLTLGLFRINESAGEIILIDGINIAKIGLHDLKFKITIIIPQDPVLV 1380
Db 1321 EKVGIVRTAGKSLTLGLFRINESAGEIILIDGINIAKIGLHDLKFKITIIIPQDPVLV 1380
Qy 1381 SGLSRMLDPSQYSDSEWNTSLLEAHLKDFVSALEPKLDHECAEGENISVCGOROLVCL 1440
Db 1381 SGLSRMLDPSQYSDSEWNTSLLEAHLKDFVSALEPKLDHECAEGENISVCGOROLVCL 1440
Qy 1441 ARALLRTKILVDEATAVADLETTDILQSTIRTOFEDCTVLITAHRLNTIMDYTRVIVL 1500
Db 1441 ARALLRTKILVDEATAVADLETTDILQSTIRTOFEDCTVLITAHRLNTIMDYTRVIVL 1500
Qy 1501 DKGEIOEYGA PSDILOQRGLFYGMADAGLV 1531
Db 1501 DKGEIOEYGA PSDILOQRGLFYGMADAGLV 1531

RESULT 13
US-08-407-207A-2
Sequence 2, Application US/08407207A
Patent No. 6063621
GENERAL INFORMATION:
APPLICANT: Deeley, Roger G.
APPLICANT: Cole, Susan P.C.
TITLE OF INVENTION: ANTIBODIES TO A MULTIDRUG RESISTANCE PROTEIN
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESSES:
ADDRESS: PARTEQ RESEARCH & DEVELOPMENT INNOVATIONS
STREET: Queen's University at Kingston
CITY: Kingston
STATE: Ontario
COUNTRY: CANADA
ZIP: K7L 3N6
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/407,207A
FILING DATE: 20-MAR-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Miernicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: 01512
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1531 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-407-207A-2

Query Match 78.4% Score 7849; DB 2; Length 1531;
 Best Local Similarity 99.9%; Pred. No. 0; Mismatches 1529; Conservative 0; Indels 2; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTVMTNSNDFPKCFONTLVWVPCFYMACPEFYLYSRH 60
 DB 1 MALRGFCSADGSDPLMDMNTVMTNSNDFPKCFONTLVWVPCFYMACPEFYLYSRH 60

QY 61 DRGIVQMTPLNKTALGFLMIVCAMDLFYSPWERSRGIFLAFVFLVSPLLGITTLLA 120
 DB 61 DRGIVQMTPLNKTALGFLMIVCAMDLFYSPWERSRGIFLAFVFLVSPLLGITTLLA 120

QY 121 TPLIQERRRGVSSGIMLTFMVALVACALAIISKIMTLKEDAOVDLPEDITFYVYFS 180
 DB 121 TPLIQERRRGVSSGIMLTFMVALVACALAIISKIMTLKEDAOVDLPEDITFYVYFS 180

QY 181 LLLIQVLVSCFSDSPLEFSETIHDNPCPESSASFLSRITFMTTGLIVRGYRPLEGSD 240
 DB 181 LLLIQVLVSCFSDSPLEFSETIHDNPCPESSASFLSRITFMTTGLIVRGYRPLEGSD 240

QY 241 LMSLNKEDTSEQVVPVLVKWKKECAKTRKQPVKVYSSKDPAPKESKVDANEVEAL 300
 DB 241 LMSLNKEDTSEQVVPVLVKWKKECAKTRKQPVKVYSSKDPAPKESKVDANEVEAL 300

QY 301 IVSPPOKEMNPSPLEKVLKTFGPFLMSFFFKALHDMFSGPQILKLLKFNVDTRAPD 360
 DB 301 IVSPPOKEMNPSPLEKVLKTFGPFLMSFFFKALHDMFSGPQILKLLKFNVDTRAPD 360

QY 361 WQGFYVLLFVTPACLOTVLVHQYFHI CFVSGMKIKTAVIGAVYRKALVITNSARKSSTV 420
 DB 361 WQGFYVLLFVTPACLOTVLVHQYFHI CFVSGMKIKTAVIGAVYRKALVITNSARKSSTV 420

QY 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVITLALYLLMLTGBSVLAGAVVWLVAVN 480
 DB 421 GEIVNLMSVDAQRFMDLATYINMISAPLOVITLALYLLMLTGBSVLAGAVVWLVAVN 480

QY 481 ANVAMKTKTYQVAMKSKDNRIKLMNEILINGIKYLKIYANELAFKDYLAIRQELKVLK 540
 DB 481 ANVAMKTKTYQVAMKSKDNRIKLMNEILINGIKYLKIYANELAFKDYLAIRQELKVLK 540

QY 541 KSAVLSAVGTFTWCTCFPLVALCTFAYVYVITDENNIIDAOAFSLALPILIRPELILP 600
 DB 541 KSAVLSAVGTFTWCTCFPLVALCTFAYVYVITDENNIIDAOAFSLALPILIRPELILP 600

QY 601 MVISSIVQASVSLKRLRIFLSHELEBDSIEREPVKDGGGNSITVYNNATFTWARSDDPT 660
 DB 601 MVISSIVQASVSLKRLRIFLSHELEBDSIEREPVKDGGGNSITVYNNATFTWARSDDPT 660

QY 720 661 LNCITTSIPREGALVAVVGVGCCGSSLLSALLAMEMDVEGHVAIKGSVAVYPOQAMQND 720
 DB 720 661 LNCITTSIPREGALVAVVGVGCCGSSLLSALLAMEMDVEGHVAIKGSVAVYPOQAMQND 720

QY 780 721 SLRENIIFGQLEBPYRSYIOACALLPDLIELPSGRTEIGEGVNLSSGQKORVSILAR 780
 DB 721 SLRENIIFGQLEBPYRSYIOACALLPDLIELPSGRTEIGEGVNLSSGQKORVSILAR 780

QY 840 781 AVYSNNDIYLPDPLSAVDAHVGHILPENYIGPKMLKNKTRILLVTHSMSYLPQVDYIIV 840
 DB 781 AVYSNNDIYLPDPLSAVDAHVGHILPENYIGPKMLKNKTRILLVTHSMSYLPQVDYIIV 840

QY 900 841 MSGGKISEMGSYOBLLARDAFAEFLTYASTEOBDAEENGYVSGPGEKAKOMENGM 900
 DB 841 MSGGKISEMGSYOBLLARDAFAEFLTYASTEOBDAEENGYVSGPGEKAKOMENGM 900

QY 960 901 LVYDSAGKOKOROLSSSSSYSGDISRHNSYALOKAKKEEFWKMEADKXQGTQVYL 960
 DB 901 LVYDSAGKOKOROLSSSSSYSGDISRHNSYALOKAKKEEFWKMEADKXQGTQVYL 960

QY 1020 961 SVYWDYKAIKGLFISLIFLFCNHNVSALASNTWLSLWTDPIVNGTQHTKVRLSVYG 1020
 DB 961 SVYWDYKAIKGLFISLIFLFCNHNVSALASNTWLSLWTDPIVNGTQHTKVRLSVYG 1020

QY 1021 ALGISOGIAVFGXSMAYSIGIILASRCLAHVDLHSLIRSMPSPFERTPSGNLVNRFSEKL 1080
 DB 1021 ALGISOGIAVFGXSMAYSIGIILASRCLAHVDLHSLIRSMPSPFERTPSGNLVNRFSEKL 1080

QY 1081 DTVDSMIPKVIKMFMSLFNVIGACIVILLATPILAIIPPLGLIYFVQRFVYASSROL 1140
 DB 1081 DTVDSMIPKVIKMFMSLFNVIGACIVILLATPILAIIPPLGLIYFVQRFVYASSROL 1140

QY 1141 KRLSVRSRSPVYSHFNETLLGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANWLA 1200
 DB 1141 KRLSVRSRSPVYSHFNETLLGVSVIRAFEEQERFIHOSDLKVDENQKAYPSIVANWLA 1200

QY 1201 VRLKESVETKEKAPMIOETAPSSWPQYGRVFRNYCYRVEDDLFVLRHINVTINGG 1320
 DB 1201 VRLKESVETKEKAPMIOETAPSSWPQYGRVFRNYCYRVEDDLFVLRHINVTINGG 1320

QY 1321 EKVIGVGTAGKSSLTGLFRINESAGEI11IDGINIAKIGLHDLRFKTI11PODPVLF 1380
 DB 1321 EKVIGVGTAGKSSLTGLFRINESAGEI11IDGINIAKIGLHDLRFKTI11PODPVLF 1380

QY 1381 SGSLRNMLDPFSGQSDSEVWTSLELAHDKDFVSALDPKLDHECAGEGNSVQOROLVCL 1440
 DB 1381 SGSLRNMLDPFSGQSDSEVWTSLELAHDKDFVSALDPKLDHECAGEGNSVQOROLVCL 1440

QY 1441 ARALLRKTILVDEATAVADLETTDLIOSTIRTOFEDCTVLTIAHLNMTIMDTRYIVL 1500
 DB 1441 ARALLRKTILVDEATAVADLETTDLIOSTIRTOFEDCTVLTIAHLNMTIMDTRYIVL 1500

QY 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531
 DB 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531

RESULT 14
 US-08-463-092B-6
 ; Sequence 6, Application US/08463092B
 ; Patent No. 5766880
 ; GENERAL INFORMATION:
 ; APPLICANT: Cole, Susan P.C.
 ; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ENCODING
 ; TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PARTIQ RESEARCH & DEVELOPMENT INNOVATIONS
 ; STREET: Queen's University at Kingston
 ; CITY: Kingston
 ; STATE: Ontario
 ; COUNTRY: CANADA
 ; ZIP: K7L 3N6
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/463,092B
 ; FILING DATE: 05-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/966,923
 ; FILING DATE: 27-OCT-1992
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/029,340
 ; FILING DATE: 8-MAR-1993
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/141,893

FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Steeg, Carol Mternicki
REGISTRATION NUMBER: 39,539
REFERENCE/DOCKET NUMBER: Q1546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (613) 545-2342
TELEFAX: (613) 545-6853
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-092B-6

Query Match 69.9%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRGFCSADGSDPLMDNVTWNTSNDPTKCFONTVLWVPCFYLMACPFYFLYLSRH 60
DB 1 MALRGFCSADGSDPLMDNVTWNTSNDPTKCFONTVLWVPCFYLMACPFYFLYLSRH 60
QY 61 DRGIQMTPLNKTALGFLMIVCMADLFYSFWRSGVIFLAPFLVSPILLGITTLA 120
DB 61 DRGIQMTPLNKTALGFLMIVCMADLFYSFWRSGVIFLAPFLVSPILLGITTLA 120
QY 121 TFLQLRRRKGVSSGIMLTFMLVALVCALILRSKIMTALKEBDAVDLFRDITFYVYS 180
DB 121 TFLQLRRRKGVSSGIMLTFMLVALVCALILRSKIMTALKEBDAVDLFRDITFYVYS 180
QY 121 TFLQLRRRKGVSSGIMLTFMLVALVCALILRSKIMTALKEBDAVDLFRDITFYVYS 180
DB 121 TFLQLRRRKGVSSGIMLTFMLVALVCALILRSKIMTALKEBDAVDLFRDITFYVYS 180
QY 181 LLLIQLVLSGSDSPLESETIHDNPNCPRESSASFLSRITFWITGLIVRGYRQPLEGSD 240
DB 181 LLLIQLVLSGSDSPLESETIHDNPNCPRESSASFLSRITFWITGLIVRGYRQPLEGSD 240
QY 241 LMSLNKEDTSQVQPVLYVNMKKECAKTRKOPVAVYVS-SKDPAPCKSSKVDANEVEEA 299
DB 241 LMSLNKEDTSQVQPVLYVNMKKECAKTRKOPVAVYVS-SKDPAPCKSSKVDANEVEEA 299
QY 241 LMSLNKEDTSQVQPVLYVNMKKECAKTRKOPVAVYVS-SKDPAPCKSSKVDANEVEEA 300
DB 241 LMSLNKEDTSQVQPVLYVNMKKECAKTRKOPVAVYVS-SKDPAPCKSSKVDANEVEEA 300
QY 300 LIVSPKQENPSPLEKVLTKTFEPFLMSFFKALHDMFSGPOLILKIFVNDTKAP 359
DB 300 LIVSPKQENPSPLEKVLTKTFEPFLMSFFKALHDMFSGPOLILKIFVNDTKAP 359
QY 301 LIVSPKQENPSPLEKVLTKTFEPFLMSFFKALHDMFSGPOLILKIFVNDTKAP 360
DB 301 LIVSPKQENPSPLEKVLTKTFEPFLMSFFKALHDMFSGPOLILKIFVNDTKAP 360
QY 360 DMGQFYFTVLLFVTAQIOTLVLAHOYFHCIVSGMKIKITAVIGAVYRKALVITNSARKSST 419
DB 360 DMGQFYFTVLLFVTAQIOTLVLAHOYFHCIVSGMKIKITAVIGAVYRKALVITNSARKSST 419
QY 361 DMGQFYFTVLLFVTAQIOTLVLAHOYFHCIVSGMKIKITAVIGAVYRKALVITNSARKSST 420
DB 361 DMGQFYFTVLLFVTAQIOTLVLAHOYFHCIVSGMKIKITAVIGAVYRKALVITNSARKSST 420
QY 420 VGEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYELMLSGPSVLAGAVVIMLWPL 479
DB 420 VGEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYELMLSGPSVLAGAVVIMLWPL 479
QY 421 VGEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYELMLSGPSVLAGAVVIMLWPL 480
DB 421 VGEIVNLSVDAQRFMDLATYINMIWSAPLOVILALYELMLSGPSVLAGAVVIMLWPL 480
QY 480 NAVVAMTKTYQVAMKSKDKRIRKLMNLIINGIKVLKLYAMELAFKDVLAIRQBELKVL 539
DB 480 NAVVAMTKTYQVAMKSKDKRIRKLMNLIINGIKVLKLYAMELAFKDVLAIRQBELKVL 539
QY 481 NAVVAMTKTYQVAMKSKDKRIRKLMNLIINGIKVLKLYAMELAFKDVLAIRQBELKVL 540
DB 481 NAVVAMTKTYQVAMKSKDKRIRKLMNLIINGIKVLKLYAMELAFKDVLAIRQBELKVL 540
QY 540 KKSAYLSAVGTFTWCTPFLVALCTFAVYVITIDENNIIIDAOTAFVSLAFNIRLPLNLT 599
DB 540 KKSAYLSAVGTFTWCTPFLVALCTFAVYVITIDENNIIIDAOTAFVSLAFNIRLPLNLT 599
QY 541 KKSAYLSAVGTFTWCTPFLVALCTFAVYVITIDENNIIIDAOTAFVSLAFNIRLPLNLT 600
DB 541 KKSAYLSAVGTFTWCTPFLVALCTFAVYVITIDENNIIIDAOTAFVSLAFNIRLPLNLT 600
QY 600 PMVSSIVQASVSLKRLRIFLSHEELBDSIERRPVKDGGGTSNITVFNATFTWASRDP 659
DB 600 PMVSSIVQASVSLKRLRIFLSHEELBDSIERRPVKDGGGTSNITVFNATFTWASRDP 659
QY 601 PMVSSIVQASVSLKRLRIFLSHEELBDSIERRSIKSGEG-NSITVQNAFTWARGBP 659
DB 601 PMVSSIVQASVSLKRLRIFLSHEELBDSIERRSIKSGEG-NSITVQNAFTWARGBP 659
QY 660 TLNGITSIPEGALVAVVGQVCGKSSLSALLAEMDKVEBHVAMIKSVAVVPOQAWION 719
DB 660 TLNGITSIPEGALVAVVGQVCGKSSLSALLAEMDKVEBHVAMIKSVAVVPOQAWION 719

QY 720 DSLRENTLFCQQLREPPYRSVIAQCALLPDLLEILPSGDRTEIGKGVNLSCGQKORVSLA 779
DB 720 DSLRENTLFCQQLREPPYRSVIAQCALLPDLLEILPSGDRTEIGKGVNLSCGQKORVSLA 779
QY 780 RAVYSNADITLFDPLSAVDAHVGKHTFENVIGKGMKXKTRILYHNSYLPQVYII 839
DB 780 RAVYSNADITLFDPLSAVDAHVGKHTFENVIGKGMKXKTRILYHNSYLPQVYII 839
QY 840 VMSGKISEMGSYOELLARDGAFAPLRTYASTEOBDAEENGTVGSPGKEAKOMENG 899
DB 840 VMSGKISEMGSYOELLARDGAFAPLRTYASTEOBDAEENGTVGSPGKEAKOMENG 899
QY 840 VMSGKISEMGSYOELLARDGAFAPLRTYASTEOBDAEENGTVGSPGKEAKOMENG 896
DB 840 VMSGKISEMGSYOELLARDGAFAPLRTYASTEOBDAEENGTVGSPGKEAKOMENG 896
QY 900 MLVTDASGKOLQRLSSSSSSGDISHHNSTAELOAKAEKFTWKLMEADKQOTGVK 959
DB 900 MLVTDASGKOLQRLSSSSSSGDISHHNSTAELOAKAEKFTWKLMEADKQOTGVK 959
QY 960 LSVYWDYMKALGLPISFLSTFLPCHNVSAALASVYMLTMDP-PYNGOEHKVLVS 1018
DB 960 LSVYWDYMKALGLPISFLSTFLPCHNVSAALASVYMLTMDP-PYNGOEHKVLVS 1018
QY 956 LSVYWDYMKALGLPISFLSTFLPCHNVSAALASVYMLTMDP-PYNGOEHKVLVS 1015
DB 956 LSVYWDYMKALGLPISFLSTFLPCHNVSAALASVYMLTMDP-PYNGOEHKVLVS 1015
QY 1019 YGALGISQGLAVFGYSMAVSIQGLASRCLHVDLHSLSPMSFFERTSGNLVNRFSK 1078
DB 1019 YGALGISQGLAVFGYSMAVSIQGLASRCLHVDLHSLSPMSFFERTSGNLVNRFSK 1078
QY 1016 YGALGISQGLAVFGYSMAVSIQGLASRCLHVDLHSLSPMSFFERTSGNLVNRFSK 1075
DB 1016 YGALGISQGLAVFGYSMAVSIQGLASRCLHVDLHSLSPMSFFERTSGNLVNRFSK 1075
QY 1079 ELDTVDSMTPEVITKMGSLFENVIGACTIVILATPILAIIPYGLIYFVQRYVASSR 1138
DB 1079 ELDTVDSMTPEVITKMGSLFENVIGACTIVILATPILAIIPYGLIYFVQRYVASSR 1138
QY 1076 ELDTVDSMTPEVITKMGSLFENVIGACTIVILATPILAIIPYGLIYFVQRYVASSR 1135
DB 1076 ELDTVDSMTPEVITKMGSLFENVIGACTIVILATPILAIIPYGLIYFVQRYVASSR 1135
QY 1139 QLRKLESVSRSPVYSHENETLIGSVIRAFEDQRFHOSDKVDENOKAYPSIVANRW 1198
DB 1139 QLRKLESVSRSPVYSHENETLIGSVIRAFEDQRFHOSDKVDENOKAYPSIVANRW 1198
QY 1136 QLRKLESVSRSPVYSHENETLIGSVIRAFEDQRFHOSDKVDENOKAYPSIVANRW 1195
DB 1136 QLRKLESVSRSPVYSHENETLIGSVIRAFEDQRFHOSDKVDENOKAYPSIVANRW 1195
QY 1199 LAVLECGVNCIVLFAVIRSHSLASGLVGLSVSYQVTTYLWMLVMSSEMETNI 1258
DB 1199 LAVLECGVNCIVLFAVIRSHSLASGLVGLSVSYQVTTYLWMLVMSSEMETNI 1258
QY 1196 LAVLECGVNCIVLFAVIRSHSLASGLVGLSVSYQVTTYLWMLVMSSEMETNI 1255
DB 1196 LAVLECGVNCIVLFAVIRSHSLASGLVGLSVSYQVTTYLWMLVMSSEMETNI 1255
QY 1259 VAVERLKEYSETEKADWQIOETAPSSWPQVGEFRNATCLARYEDLDVLAIRHINVTIN 1318
DB 1259 VAVERLKEYSETEKADWQIOETAPSSWPQVGEFRNATCLARYEDLDVLAIRHINVTIN 1318
QY 1319 GGEKVIGVTRGAKSSLTGLPRINSAGEIITIDINIAKIGLHDLRPFKTIITIPDDPV 1378
DB 1319 GGEKVIGVTRGAKSSLTGLPRINSAGEIITIDINIAKIGLHDLRPFKTIITIPDDPV 1378
QY 1316 GGEKVIGVTRGAKSSLTGLPRINSAGEIITIDINIAKIGLHDLRPFKTIITIPDDPV 1375
DB 1316 GGEKVIGVTRGAKSSLTGLPRINSAGEIITIDINIAKIGLHDLRPFKTIITIPDDPV 1375
QY 1379 LFGSLRNNLDPFQYDEEVTSLFLAHLKDFFVSALPDKLDHECAGENLTVGQROLV 1438
DB 1379 LFGSLRNNLDPFQYDEEVTSLFLAHLKDFFVSALPDKLDHECAGENLTVGQROLV 1438
QY 1376 LFGSLRNNLDPFQYDEEVTSLFLAHLKDFFVSALPDKLDHECAGENLTVGQROLV 1435
DB 1376 LFGSLRNNLDPFQYDEEVTSLFLAHLKDFFVSALPDKLDHECAGENLTVGQROLV 1435
QY 1439 CLARALLRKTILVLDATAVADLETDLIQSTIRIQFEDCTVLTIAHRLMTINDYRVI 1498
DB 1439 CLARALLRKTILVLDATAVADLETDLIQSTIRIQFEDCTVLTIAHRLMTINDYRVI 1498
QY 1436 CLARALLRKTILVLDATAVADLETDLIQSTIRIQFEDCTVLTIAHRLMTINDYRVI 1495
DB 1436 CLARALLRKTILVLDATAVADLETDLIQSTIRIQFEDCTVLTIAHRLMTINDYRVI 1495
QY 1499 VLDKGEIOEYGABSDLLQORGLFYSMAKDAGLV 1531
DB 1499 VLDKGEIOEYGABSDLLQORGLFYSMAKDAGLV 1531
QY 1496 VLDKGEIOEYGABSDLLQORGLFYSMAKDAGLV 1528
DB 1496 VLDKGEIOEYGABSDLLQORGLFYSMAKDAGLV 1528

RESULT 15
US-08-462-109A-6
Sequence 6, Application US/08462109A
Patent No. 5882875
GENERAL INFORMATION:
APPLICANT: Cole, Susan P. C.
APPLICANT: Dealey, Roger G.
TITLE OF INVENTION: METHODS FOR IDENTIFYING
NUMBER OF SEQUENCES: 6
TITLE OF INVENTION: MULTIDRUG RESISTANT TUMOR CELLS
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA

ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/462,109A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/966,923
FILING DATE: 27-OCT-1992
APPLICATION NUMBER: 08/029,340
FILING DATE: 8-MAR-1993
APPLICATION NUMBER: 08/141,893
FILING DATE: 26-OCT-1993
APPLICATION NUMBER: 08/407,207
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Decont, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: PQ1-002CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-462-109A-6

Query Match 69.9%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRGFSADGSDPLDMNMTNTNSNDFKCFONTYLVWPCYYLWACPPFFLYSRH 60
DB 1 MAIRSFSDGSDPLDMNMTNTNSNDFKCFONTYLVWPCYYLWACPPFFLYSRH 60
QY 61 DRGVIQWTPINKTKTALGFLMIVCMADLFYSFWERSRGIFLAPVLPVSPFLIGITLLA 120
DB 61 DRGVIQWTPINKTKTALGFLMIVCMADLFYSFWERSRGIFLAPVLPVSPFLIGITLLA 120
QY 121 TFLIQERRRGVSSGIMLTFMVALVCAALILRSKIMTALKEDAQYDLFRDITFYVYS 180
DB 121 TFLIQERRRGVSSGIMLTFMVALVCAALILRSKIMTALKEDAQYDLFRDITFYVYS 180
QY 181 LLLIQVLSCFSDRSPLFSETIHDNCPSSSASFSLRTTFWMTGLIVRGYROPLEGS 240
DB 181 LLLIQVLSCFSDRSPLFSETIHDNCPSSSASFSLRTTFWMTGLIVRGYROPLEGS 240
QY 241 LMSLNKEDTSEOVYVPLVKNMKKCAKTRKOPVNVYS-SKDPAPCKESSKVAANEVEA 299
DB 241 LMSLNKEDTSEOVYVPLVKNMKKCAKTRKOPVNVYS-SKDPAPCKESSKVAANEVEA 299
QY 300 LLYKSPCKEWNPSLFKLYKTFGPFYPLMSFFKAIHDLMMFSGPQILKLIKFNVDTKAP 359
DB 300 LLYKSPCKEWNPSLFKLYKTFGPFYPLMSFFKAIHDLMMFSGPQILKLIKFNVDTKAP 359
QY 360 DMGCFYTVLLFTYACIQTLVHQYFHIQYSGMRITAVIYGAAYRKALVITNSARKST 419
DB 360 DMGCFYTVLLFTYACIQTLVHQYFHIQYSGMRITAVIYGAAYRKALVITNSARKST 419
QY 420 VGEIVNLSYDAORFMDLATYINMWSAPVOVITATLWLNIGPSYLAGVAVWVWVAV 479
DB 420 VGEIVNLSYDAORFMDLATYINMWSAPVOVITATLWLNIGPSYLAGVAVWVWVAV 479
QY 480 NAVVAMKTKTYQVAHMSKDNRIKLMNEILNGIKVLKLYVAMELAFKDKVLAIRQEBELKVL 539
DB 480 NAVVAMKTKTYQVAHMSKDNRIKLMNEILNGIKVLKLYVAMELAFKDKVLAIRQEBELKVL 539

QY 540 KKSAYLSAVGFFTWCTPFLVALCTPAVYVTTIDENNILDAQAFVSLATFNILRPPLNLT 599
DB 541 KKSAYLSAVGFFTWCTPFLVALCTPAVYVTTIDENNILDAQAFVSLATFNILRPPLNLT 600
QY 600 PMVSIIVQASVSLKRLRIFLSHELEPDSIERRPVDGGGTSITVNNATFTWASDP 659
DB 601 PMVSIIVQASVSLKRLRIFLSHELEPDSIERRPVDGGGTSITVNNATFTWASDP 659
QY 660 TLNGITSPGALVAVVGVCGKSSLSALLAEMDKVEGHVAKGSVAVVPOQAMON 719
DB 660 TLNGITSPGALVAVVGVCGKSSLSALLAEMDKVEGHVAKGSVAVVPOQAMON 719
QY 720 DSIARENLPFGQLEPPYRSVIOACALLPDELTPSGDRPTEIGKGVNLSSGQQRVSLA 779
DB 720 DSIARENLPFGQLEPPYRSVIOACALLPDELTPSGDRPTEIGKGVNLSSGQQRVSLA 779
QY 780 RAVYSNADITLFDPLSAVDAHVGKHI FENVIGPKMLKXKTRILVTHSMNYPQVYII 839
DB 780 RAVYSNADITLFDPLSAVDAHVGKHI FENVIGPKMLKXKTRILVTHSMNYPQVYII 839
QY 840 VMSGKISEMGSYQELLARDGAPAFPLRTYASTEOEDAEENGVTGVSGPKAKQWENG 899
DB 840 VMSGKISEMGSYQELLARDGAPAFPLRTYASTEOEDAEENGVTGVSGPKAKQWENG 899
QY 900 MLVTDAGKQOLQOLSSSSSYSGDISRHHNSTAELQAEAKKERTKLMADKAQOTGVK 959
DB 900 MLVTDAGKQOLQOLSSSSSYSGDISRHHNSTAELQAEAKKERTKLMADKAQOTGVK 959
QY 960 LSVYMDYKAIIGLFIPLSLIFLPMCNHVSALASVYMLSTWTD-PYNGTOEHKRVLSV 1018
DB 960 LSVYMDYKAIIGLFIPLSLIFLPMCNHVSALASVYMLSTWTD-PYNGTOEHKRVLSV 1018
QY 1019 YGALGISQGIATVGYMAVSIIGGILASRCHAVDLNHSILRSPMSFFERTSPGNLVNRFK 1078
DB 1019 YGALGISQGIATVGYMAVSIIGGILASRCHAVDLNHSILRSPMSFFERTSPGNLVNRFK 1078
QY 1079 ELDTVDSMTEPVKIMFPGSLFNVIYACIVILLATPIAIIIIIPGLGIFVYQRYVASR 1138
DB 1079 ELDTVDSMTEPVKIMFPGSLFNVIYACIVILLATPIAIIIIIPGLGIFVYQRYVASR 1138
QY 1139 QLRKLESVSRSPVYSHNETLLGVSVARPEBERFHOSDLKVDENQKAYPSIVANRW 1198
DB 1139 QLRKLESVSRSPVYSHNETLLGVSVARPEBERFHOSDLKVDENQKAYPSIVANRW 1198
QY 1199 LAVRLCEVNCIYLFALFAVISRHSLSAGLVGLSVYSIQVTTYLNMLVMSSEMETNI 1258
DB 1199 LAVRLCEVNCIYLFALFAVISRHSLSAGLVGLSVYSIQVTTYLNMLVMSSEMETNI 1258
QY 1259 VAVRLKESYSETEKAPWQIOETAPPSWPOVRVEFRNYCLARYREDDLVLRHINVTIN 1318
DB 1259 VAVRLKESYSETEKAPWQIOETAPPSWPOVRVEFRNYCLARYREDDLVLRHINVTIN 1318
QY 1319 GGEKVGIVGRTGAKSLTGLFRINESAGEIIIDGINIAKIGLHDLRKITIIIPDDV 1378
DB 1319 GGEKVGIVGRTGAKSLTGLFRINESAGEIIIDGINIAKIGLHDLRKITIIIPDDV 1378
QY 1379 LFSGSLRMNLDPPSOYSDSEBWTSLBLAHLKDVVSALPDLKDEHCAEGENLTVGQROLY 1438
DB 1379 LFSGSLRMNLDPPSOYSDSEBWTSLBLAHLKDVVSALPDLKDEHCAEGENLTVGQROLY 1438
QY 1439 CLARALLRKTKIIVLDEBATVAVDLEFDNLIQSTIRPOFEDCTVLTIAHRLNTMDYTRVI 1498
DB 1439 CLARALLRKTKIIVLDEBATVAVDLEFDNLIQSTIRPOFEDCTVLTIAHRLNTMDYTRVI 1498
QY 1499 VLDKGEIOEYGAPSDLDQORGLFYMAKXAGIV 1531
DB 1499 VLDKGEIOEYGAPSDLDQORGLFYMAKXAGIV 1531

Search completed: December 15, 2005, 15:50:04
Job time : 52.4867 secs

This Page Blank (copy)

QY 361 MGGFYTVLLFVTAQCTLVLAHQYFHICFVSGMBIKTAIVGAVYRKALVITNSARKSTV 420
Db 361 MGGFYTVLLFVTAQCTLVLAHQYFHICFVSGMBIKTAIVGAVYRKALVITNSARKSTV 420
QY 421 GEIYNLMSVDAQRMDLATTYINMTWSAPLOYITLALYLMNLGSPVLAGAVNMLAMPVN 480
Db 421 GEIYNLMSVDAQRMDLATTYINMTWSAPLOYITLALYLMNLGSPVLAGAVNMLAMPVN 480
QY 481 AVAMMKTKTQVAMMKSKDNRIKLMNEILNGIKYUKLYAMELAPDKYLAIOBELKYLK 540
Db 481 AVAMMKTKTQVAMMKSKDNRIKLMNEILNGIKYUKLYAMELAPDKYLAIOBELKYLK 540
QY 541 KSAIYASVGTFTWCTPFLVALCTPAVYVTTIDENNITLDAQTAFFVSLAFNLIREPLNLP 600
Db 541 KSAIYASVGTFTWCTPFLVALCTPAVYVTTIDENNITLDAQTAFFVSLAFNLIREPLNLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPYKOGGCTNSITVRNATTTMARSDPT 660
Db 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPYKOGGCTNSITVRNATTTMARSDPT 660
QY 661 LINGITFSPREGALVAVVGVQVCGKSSLLSALLAEMDKVGHVAIKGSVAVYPOQAMIOND 720
Db 661 LINGITFSPREGALVAVVGVQVCGKSSLLSALLAEMDKVGHVAIKGSVAVYPOQAMIOND 720
QY 721 SLRENILFGCOLLEBPYRSVIOACALPDLLEILPSGDRTEIGEKGVNLSGGOKORVSLAR 780
Db 721 SLRENILFGCOLLEBPYRSVIOACALPDLLEILPSGDRTEIGEKGVNLSGGOKORVSLAR 780
QY 781 AVYENADITFLPDDLSAVDAVGHKIFENVIGPKMKLKNKTRILVTHSMASLPOVDVITV 840
Db 781 AVYENADITFLPDDLSAVDAVGHKIFENVIGPKMKLKNKTRILVTHSMASLPOVDVITV 840
QY 841 MSGGKISMGSYOELLARDGAPAEFLRTYASTBOQDABENGCVTSVPGKBAKOMENGM 900
Db 841 MSGGKISMGSYOELLARDGAPAEFLRTYASTBOQDABENGCVTSVPGKBAKOMENGM 900
QY 901 LVTDASAGKOLOROLSSSSSSSGDISRHHNSTAELOKAKKEETWKLMEADKAQGVYKL 960
Db 901 LVTDASAGKOLOROLSSSSSSSGDISRHHNSTAELOKAKKEETWKLMEADKAQGVYKL 960
QY 961 SVYWDYMKALGIFLSIFSLIFLPMCHVSAALSNVWLSMTDDPIYNGIOEHTKRLSYG 1020
Db 961 SVYWDYMKALGIFLSIFSLIFLPMCHVSAALSNVWLSMTDDPIYNGIOEHTKRLSYG 1020
QY 1021 ALGISOGIAVFGYSAVSIIGILASRCHVDLHLSILSPSPFERTSGNLYNRFSEKL 1080
Db 1021 ALGISOGIAVFGYSAVSIIGILASRCHVDLHLSILSPSPFERTSGNLYNRFSEKL 1080
QY 1081 DTVDSMIPEVIKMFMSLFFNYIACIVILLATPIAIIIPPLGLIYFFVQRFYVYASSHQL 1140
Db 1081 DTVDSMIPEVIKMFMSLFFNYIACIVILLATPIAIIIPPLGLIYFFVQRFYVYASSHQL 1140
QY 1141 KRLSESVSSPYVSHNETLLGVSVIRAFEBQERFIHQSDLKVDENOKAYYSIVANRWLA 1200
Db 1141 KRLSESVSSPYVSHNETLLGVSVIRAFEBQERFIHQSDLKVDENOKAYYSIVANRWLA 1200
QY 1201 VRLCEVGCIVLFAALFAVIRHSLSAGLNGLSVYSIQVTTYLNMVLRMSSEMTNIVA 1260
Db 1201 VRLCEVGCIVLFAALFAVIRHSLSAGLNGLSVYSIQVTTYLNMVLRMSSEMTNIVA 1260
QY 1261 VERLEKEYETEKEAPWQIOETAPSPSPQVGVREPRNCLARYEDLDVLRHINVTINGG 1320
Db 1261 VERLEKEYETEKEAPWQIOETAPSPSPQVGVREPRNCLARYEDLDVLRHINVTINGG 1320
QY 1321 EKVGVIGRTGAKSSLLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKRTITIIPODPVLF 1380
Db 1321 EKVGVIGRTGAKSSLLTGLFRINESAEGEIIIDGINIAKIGLHDLRFKRTITIIPODPVLF 1380
QY 1381 SCSLRMLNDPFSQYDEDEWMTSLAHLKDPVSLPDLXDHCAAGGNLSVGGQOLYCL 1440
Db 1381 SCSLRMLNDPFSQYDEDEWMTSLAHLKDPVSLPDLXDHCAAGGNLSVGGQOLYCL 1440
QY 1440 SCSLRMLNDPFSQYDEDEWMTSLAHLKDPVSLPDLXDHCAAGGNLSVGGQOLYCL 1440

QY 1441 APALARKTKLVLDEATAVNDLETDLIQSTIRTOFEDCVLTIARHLNTIMDYTRIVL 1500
Db 1441 APALARKTKLVLDEATAVNDLETDLIQSTIRTOFEDCVLTIARHLNTIMDYTRIVL 1500
QY 1501 DKGEIOEYGA PSDLQORGFYSMAKADAGLVGGGGGMSLRKGIIPREYVLTREAPAE 1560
Db 1501 DKGEIOEYGA PSDLQORGFYSMAKADAGLVGGGGGMSLRKGIIPREYVLTREAPAE 1560
QY 1561 PRYTRERRARRFVSKKNCNVANHKNIREQGRFLQDVFETTVLXKMPHTLLIFTMSFLCSW 1620
Db 1561 PRYTRERRARRFVSKKNCNVANHKNIREQGRFLQDVFETTVLXKMPHTLLIFTMSFLCSW 1620
QY 1621 LLEFAMVWMLIAFAHGDLYAYMEKGITDLAPYPYDVPDYAGBNVPCVTSIHSFSAFLF 1680
Db 1621 LLEFAMVWMLIAFAHGDLYAYMEKGITDLAPYPYDVPDYAGBNVPCVTSIHSFSAFLF 1680
QY 1681 SIEVQVITIGFGRNVTECPALILILVQNI VGLMNAIMLGCIFMKTQAHRAEFLIF 1740
Db 1681 SIEVQVITIGFGRNVTECPALILILVQNI VGLMNAIMLGCIFMKTQAHRAEFLIF 1740
QY 1741 SKHAVITLRHGRLCFMLRVGDLRKSMTISATIMQVVRKTTSPGEVVPPLHQVDIPMENG 1800
Db 1741 SKHAVITLRHGRLCFMLRVGDLRKSMTISATIMQVVRKTTSPGEVVPPLHQVDIPMENG 1800
QY 1801 VGGNGIFLVA PLIYHVIDSNSPLYDLAPSDLHHQDLLEIIVILEGVETTGITTOARTS 1860
Db 1801 VGGNGIFLVA PLIYHVIDSNSPLYDLAPSDLHHQDLLEIIVILEGVETTGITTOARTS 1860
QY 1861 YLADEIIMQGRFVPIVAEEDGRYSVDYSKFGNTIKVPTPLCTARQLEDSDLALTLAS 1920
Db 1861 YLADEIIMQGRFVPIVAEEDGRYSVDYSKFGNTIKVPTPLCTARQLEDSDLALTLAS 1920
QY 1921 SRGPLRRSVAVAKAKPFISIPDLS 1947
Db 1921 SRGPLRRSVAVAKAKPFISIPDLS 1947

RESULT 2
US-10-665-283-1
; Sequence 1, Application US/10665283
; Publication No. US2005063989A1
; GENERAL INFORMATION:
; APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
; APPLICANT: DERAND, Renaud
; APPLICANT: GARCIA, Elisabeth
; APPLICANT: PROST, Anne-Lise
; APPLICANT: REVILLAUD, Jean
; APPLICANT: VIVAUDOU, Michel
; TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF
; FILE REFERENCE: P263US104
; CURRENT APPLICATION NUMBER: US/10/665,283
; CURRENT FILING DATE: 2003-09-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 1927
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-665-283-1

Query Match 98.6%; Score 9873; DB 5; Length 1927;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1927; Conservative 0; Mismatches 0; Indels 20; Gaps 2;
QY 1 MALRGFCGADSDPLMDMNTWTNNTSNDFTCFQNTVLYVWPCTYIACFPFYLYLSRH 60
Db 1 MALRGFCGADSDPLMDMNTWTNNTSNDFTCFQNTVLYVWPCTYIACFPFYLYLSRH 60
QY 61 DRGIQMTPLNKKTALGFLIMIVCMADLFYSFMRSGRIFLAPVFLVSPFLITTLIA 120
Db 61 DRGIQMTPLNKKTALGFLIMIVCMADLFYSFMRSGRIFLAPVFLVSPFLITTLIA 120
QY 121 TFLIQLERRKGVQSSGIMLTFMLVALVCAALIRSKIMTALKEADAQVLDLFRDITFYVYS 180

Db 121 TFLIQLRRKQVSSGIMLTFMVLVALCALAIIRKSKIMTALKEDAQVDFPDIITFYVFS 180
 Qy 181 LLLIQLVLSCDSBSPFSEITIHDPNCPRESSASFLSITFTWMTGLIVRGYRPLEGSD 240
 Db 181 LLLIQLVLSCDSBSPFSEITIHDPNCPRESSASFLSITFTWMTGLIVRGYRPLEGSD 240
 Qy 241 LMSLNKEDTSEQVAVLVKMKKECAKTRKQPVVVVSSKDPAPKSSKTDABEVEAL 300
 Db 241 LMSLNKEDTSEQVAVLVKMKKECAKTRKQPVVVVSSKDPAPKSSKTDABEVEAL 300
 Qy 301 IVKSPQKEMNPSLFKVLKTFGPFILMSFFPKAIHDLMMFSGPOLIKLILFVNDTYAPD 360
 Db 301 IVKSPQKEMNPSLFKVLKTFGPFILMSFFPKAIHDLMMFSGPOLIKLILFVNDTYAPD 360
 Qy 361 WQGFYFVVLVFTVCLQTLVLHQYFHCIFVSGMRKITAIVGAVTRKALVITNSARKSSTV 420
 Db 361 WQGFYFVVLVFTVCLQTLVLHQYFHCIFVSGMRKITAIVGAVTRKALVITNSARKSSTV 420
 Qy 421 GEIVNLMSVDAQRFMDLATYINMTMSAPIOVILALYILMLNGPSVLGAVAVMTLMPVN 480
 Db 421 GEIVNLMSVDAQRFMDLATYINMTMSAPIOVILALYILMLNGPSVLGAVAVMTLMPVN 480
 Qy 481 AVNAAKTKTYQVAHMKSKDNRIKLMNEILNGIKVILKYAMELAFKDKVLAIROBELKVLK 540
 Db 481 AVNAAKTKTYQVAHMKSKDNRIKLMNEILNGIKVILKYAMELAFKDKVLAIROBELKVLK 540
 Qy 541 KSAVLASVGTFTWCTPFLVALCTPFAVYVITIDENNILDAQTAFLVALFNILRPLNLP 600
 Db 541 KSAVLASVGTFTWCTPFLVALCTPFAVYVITIDENNILDAQTAFLVALFNILRPLNLP 600
 Qy 601 MVLSSIVQASVSLKRLIFLISHELEBDSIERRPVQGGGNGSTTVNATFTARSPPT 660
 Db 601 MVLSSIVQASVSLKRLIFLISHELEBDSIERRPVQGGGNGSTTVNATFTARSPPT 660
 Qy 661 LMGITFESIPGALVAVVGVCGCKSSLALAEEMDKVEGHVAKGSVAYVPOQAWIOND 720
 Db 661 LMGITFESIPGALVAVVGVCGCKSSLALAEEMDKVEGHVAKGSVAYVPOQAWIOND 720
 Qy 721 SLRENILFGCQLEBPPYRSVIAOCALLPDLIELPSGRTEIGEKNVLSGGOKRVSILAR 780
 Db 721 SLRENILFGCQLEBPPYRSVIAOCALLPDLIELPSGRTEIGEKNVLSGGOKRVSILAR 780
 Qy 781 AAVSNADIYLPDDELAVDAVGHKIFENYIGPKMKIKNKRILIVTMSMSTLPDVDTIIV 840
 Db 781 AAVSNADIYLPDDELAVDAVGHKIFENYIGPKMKIKNKRILIVTMSMSTLPDVDTIIV 840
 Qy 841 MSGGKISEMGSYOELARDGAFAPFLRTYASTBOEODAEENGTVGSGPKAEAKOMENGM 900
 Db 841 MSGGKISEMGSYOELARDGAFAPFLRTYASTBOEODAEENGTVGSGPKAEAKOMENGM 900
 Qy 901 LVTDASGKOLOROLSSSSSYSGDISRRHNSTAELOKAEKKEETWKLMEADKAQTQVXL 960
 Db 901 LVTDASGKOLOROLSSSSSYSGDISRRHNSTAELOKAEKKEETWKLMEADKAQTQVXL 960
 Qy 961 SYVWDYKAKIAGLFISELIFLPMCNHVSALASNTWLSLMTDDPVLNGTOEHTKRLSVYG 1020
 Db 961 SYVWDYKAKIAGLFISELIFLPMCNHVSALASNTWLSLMTDDPVLNGTOEHTKRLSVYG 1020
 Qy 1021 ALGISOGIAVFGYSMAVISGIIASRCLHYDLHSILRSPMSFPERIPSGNLVNRFSKEL 1080
 Db 1021 ALGISOGIAVFGYSMAVISGIIASRCLHYDLHSILRSPMSFPERIPSGNLVNRFSKEL 1080
 Qy 1081 DTYDMSMPBEVIAKMFMSGLFNVIAGACIYIILATPIAIIIPPLGILYFFVGRFYAASSROL 1140
 Db 1081 DTYDMSMPBEVIAKMFMSGLFNVIAGACIYIILATPIAIIIPPLGILYFFVGRFYAASSROL 1140
 Qy 1141 KRLSVSRPVSYSFHNFTLIGSVIRAFEOERFIHOSDLKVDENQKAYVPSIYANRWLA 1200
 Db 1141 KRLSVSRPVSYSFHNFTLIGSVIRAFEOERFIHOSDLKVDENQKAYVPSIYANRWLA 1200
 Qy 1201 VRLCEVNCIIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYANMLVRMSSEMETNIIVA 1260
 Db 1201 VRLCEVNCIIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYANMLVRMSSEMETNIIVA 1260

Db 1201 VRLCEVNCIIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYANMLVRMSSEMETNIIVA 1260
 Qy 1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVFRNYCIRYRDDLPVLRHINVTINGG 1320
 Db 1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVFRNYCIRYRDDLPVLRHINVTINGG 1320
 Qy 1321 EKVGIVRTGAGKSLTLGLFRINESAGEIILIDGINIAKIGLHDLFPKTIIPQDPVLF 1380
 Db 1321 EKVGIVRTGAGKSLTLGLFRINESAGEIILIDGINIAKIGLHDLFPKTIIPQDPVLF 1380
 Qy 1381 SGLSRMLDPPSQSDEBVTSLBLAHKDFVSALPDKLDHECAGEENLSVGOROLVCL 1440
 Db 1381 SGLSRMLDPPSQSDEBVTSLBLAHKDFVSALPDKLDHECAGEENLSVGOROLVCL 1440
 Qy 1441 ARALIRTKTILVDEATAVADLETDLIIOSTIRQFEDCTVLTIAHRLNTIMDYTRYIVL 1500
 Db 1441 ARALIRTKTILVDEATAVADLETDLIIOSTIRQFEDCTVLTIAHRLNTIMDYTRYIVL 1500
 Qy 1501 DKGEIOEYGA PSDILOQRGLFYSAKADAGVGGGGGMLSRKGIIPBEVYLTRLAEPDAE 1560
 Db 1501 DKGEIOEYGA PSDILOQRGLFYSAKADAGVGGGGGMLSRKGIIPBEVYLTRLAEPDAE 1560
 Qy 1561 PRYTRERRARFVSKGNCNVAHKNIREQGRFLQDVFETTVLDKMPHTLILFTMSFLCSW 1620
 Db 1561 PRYTRERRARFVSKGNCNVAHKNIREQGRFLQDVFETTVLDKMPHTLILFTMSFLCSW 1620
 Qy 1621 LLEPAMVWMLIAFAHGDLYAEMKGITDLAPPYDVPDYAGEGTVNVCVTSIHSFSSAPLF 1680
 Db 1621 LLEPAMVWMLIAFAHGDLYAEMKGITDLAPPYDVPDYAGEGTVNVCVTSIHSFSSAPLF 1680
 Qy 1681 SIEVOVTIGRGMVTECEPLATILILVQNVIGIMINAMIGCIIPMTAQAHRAPAEFLIF 1740
 Db 1681 SIEVOVTIGRGMVTECEPLATILILVQNVIGIMINAMIGCIIPMTAQAHRAPAEFLIF 1740
 Qy 1741 SKHAVITLRHGRICFMLRVBDLRKSMIISATIHQVVRKTSPEGEVAPLHOVDI PMENG 1800
 Db 1741 SKHAVITLRHGRICFMLRVBDLRKSMIISATIHQVVRKTSPEGEVAPLHOVDI PMENG 1800
 Qy 1801 VGGNGIFLVAPLIIYHVIDSNSPLYDLAPSDLHHHODEIIVILGEVETTGITTOARTS 1860
 Db 1801 VGGNGIFLVAPLIIYHVIDSNSPLYDLAPSDLHHHODEIIVILGEVETTGITTOARTS 1860
 Qy 1861 YLADELIMGQRFVPIVAEEGRYSVYSKRGNTIKVPTPLCTARQLEDSDSLDALTLAS 1920
 Db 1861 YLADELIMGQRFVPIVAEEGRYSVYSKRGNTIKVPTPLCTARQLEDSDSLDALTLAS 1920
 Qy 1921 SRGPLRKRSVAVAKAKPKFSISPDLS 1947
 Db 1901 SRGPLRKRSVAVAKAKPKFSISPDLS 1927

RESULT 3
 US-10-665-283-6
 ; Sequence 6, Application US/10665283
 ; Publication No. US20050063989A1
 ; GENERAL INFORMATION:
 ; APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
 ; APPLICANT: DERAND, Renaud
 ; APPLICANT: GARCIA, Elisabeth
 ; APPLICANT: PROST, Anne-lise
 ; APPLICANT: REVILLIOD, Jean
 ; APPLICANT: VIVADOU, Michel
 ; TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF
 ; FILE REFERENCE: F263US104
 ; CURRENT APPLICATION NUMBER: US/10/665,283
 ; NUMBER OF SEQ. ID NOS: 25
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 1927
 ; TYPE: PRK
 ; ORGANISM: Homo sapiens
 US-10-665-283-6

Query Match 98.5%; Score 9661; DB 5; Length 1927;
 Best Local Similarity 98.9%; Pred. No. 0;
 Matches 1925; Conservative 0; Mismatches 2; Indels 20; Gaps 2;

```

Qy 1 MALRGFSSAGSDLDLMDMNTWNTSNPDTFCQNTVLVWPCRYLWACFFPYLYLSRH 60
Db 1 MALRGFSSAGSDLDLMDMNTWNTSNPDTFCQNTVLVWPCRYLWACFFPYLYLSRH 60
Qy 61 DRGTYQMTPLNKTALGFLIMIVCMADLFYSFWEBSRGIFLAPFLVSPILLGITLLA 120
Db 61 DRGTYQMTPLNKTALGFLIMIVCMADLFYSFWEBSRGIFLAPFLVSPILLGITLLA 120
Qy 121 TFLIQLERRKGVSSGIMLTFWLVACALILRSKIMTALKEDAQVDLDFDITFYVYS 180
Db 121 TFLIQLERRKGVSSGIMLTFWLVACALILRSKIMTALKEDAQVDLDFDITFYVYS 180
Qy 181 LLILQLVLSCFSDSPFLSETIHDNPNCPRESSASFLSITFWMTGLIVRGYRQPLESD 240
Db 181 LLILQLVLSCFSDSPFLSETIHDNPNCPRESSASFLSITFWMTGLIVRGYRQPLESD 240
Qy 241 LMSLNKEDTSEQVVPVLVKNMKECAKTRKQPVKVYSSKDPAPKESKVDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLVKNMKECAKTRKQPVKVYSSKDPAPKESKVDANEVEAL 300
Qy 301 IVKSPOKEMNPSLFKVLKTFGPYFLMSFFPKAIDHLMFSGPOLIKLIFVNDTRAPD 360
Db 301 IVKSPOKEMNPSLFKVLKTFGPYFLMSFFPKAIDHLMFSGPOLIKLIFVNDTRAPD 360
Qy 361 WQGFYVTLFVTCLOTLVHOYFHCFSVSGMRIKTAVIGAVRKALVITNSARKSTV 420
Db 361 WQGFYVTLFVTCLOTLVHOYFHCFSVSGMRIKTAVIGAVRKALVITNSARKSTV 420
Qy 421 GEIYNLMSVDAQRFMDLATYINMIWSAPLOVILALYLLMLNMGPSVLGAVAMVLMEVN 480
Db 421 GEIYNLMSVDAQRFMDLATYINMIWSAPLOVILALYLLMLNMGPSVLGAVAMVLMEVN 480
Qy 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAMEIAPKDYLAIROBELKYLK 540
Db 481 AVMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKLYAMEIAPKDYLAIROBELKYLK 540
Qy 541 KSAFLSAGFTFWCTPFLVALCTPAVYVYIDENNIIDAQTAFLYSALFNILRPPLNLP 600
Db 541 KSAFLSAGFTFWCTPFLVALCTPAVYVYIDENNIIDAQTAFLYSALFNILRPPLNLP 600
Qy 601 MVISSIVQASYSKRLRIFLSHEBELPDSIRRPVKDGGTNSITVRNATFTMARSDPT 660
Db 601 MVISSIVQASYSKRLRIFLSHEBELPDSIRRPVKDGGTNSITVRNATFTMARSDPT 660
Qy 661 LMGITFSIPREGALVAVVGVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAMIQND 720
Db 661 LMGITFSIPREGALVAVVGVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAMIQND 720
Qy 721 SLRENILFQCGLLEBPYRSVTOACALLPDLLEILPSGDRTEIGEKVNLSSGQOKRVSLAR 780
Db 721 SLRENILFQCGLLEBPYRSVTOACALLPDLLEILPSGDRTEIGEKVNLSSGQOKRVSLAR 780
Qy 781 AVYENADIYLFDDPLSAVDAAVGHKIFENYIGPKKMLKNKTRILVTHSMSTLPQVDVIV 840
Db 781 AVYENADIYLFDDPLSAVDAAVGHKIFENYIGPKKMLKNKTRILVTHSMSTLPQVDVIV 840
Qy 841 MSGGKISEMGSYOELLARDGAFABFLRTYASTBOQDAEENGVTGVSFGKEAKOMENGM 900
Db 841 MSGGKISEMGSYOELLARDGAFABFLRTYASTBOQDAEENGVTGVSFGKEAKOMENGM 900
Qy 901 LVTHSAGQOLOROUSSSSSYSGDISRHHNSTAELOKAKKEBTYKLMLEADKAQGTQYKL 960
Db 901 LVTHSAGQOLOROUSSSSSYSGDISRHHNSTAELOKAKKEBTYKLMLEADKAQGTQYKL 960
Qy 961 SVYDYMKAIGLFTSFLSIFLPMCHVASALSNVWLSLMTDPTVNGOETHKVALSYVG 1020
Db 961 SVYDYMKAIGLFTSFLSIFLPMCHVASALSNVWLSLMTDPTVNGOETHKVALSYVG 1020

```

```

Qy 1021 ALGISQIAVFGYSMAVSIIGIILASRCLHYDLHSIIRSPSPFPERPDSGNLVNREKEL 1080
Db 1021 ALGISQIAVFGYSMAVSIIGIILASRCLHYDLHSIIRSPSPFPERPDSGNLVNREKEL 1080
Qy 1081 DTVDSMIPEVIKMMGSLFNIVIGACIVILATPIAIIIPPLIGIYFVFORFYASSROL 1140
Db 1081 DTVDSMIPEVIKMMGSLFNIVIGACIVILATPIAIIIPPLIGIYFVFORFYASSROL 1140
Qy 1141 KRLESVRSRPSYSHFNELLGVSIVIRAFEEQERFIHQSDLKVDENOKAYPSIVANWMLA 1200
Db 1141 KRLESVRSRPSYSHFNELLGVSIVIRAFEEQERFIHQSDLKVDENOKAYPSIVANWMLA 1200
Qy 1201 VRLCEVNCIVLPAALPAVISRHSLSAGLVLSYSYSLOTTYVNTLVRRSMSEMTIVA 1260
Db 1201 VRLCEVNCIVLPAALPAVISRHSLSAGLVLSYSYSLOTTYVNTLVRRSMSEMTIVA 1260
Qy 1261 VERLKEYSETEKAPMGOIQTAPPSSMPQVRVEFRNYCLRYRBDLPVLRIHIVTTINGG 1320
Db 1261 VERLKEYSETEKAPMGOIQTAPPSSMPQVRVEFRNYCLRYRBDLPVLRIHIVTTINGG 1320
Qy 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
Db 1321 EKVGIVGRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
Qy 1381 SSGIRKMLDPSSQYSDDEEWTSLLEALIXDFVSLPKDLHECAGEGEMSVGQROLVCL 1440
Db 1381 SSGIRKMLDPSSQYSDDEEWTSLLEALIXDFVSLPKDLHECAGEGEMSVGQROLVCL 1440
Qy 1441 ABALIRKTKILVDEAFAVADLETDDLIOSTIRPOFEDCVLTATHLANTMDTRYIVL 1500
Db 1441 ABALIRKTKILVDEAFAVADLETDDLIOSTIRPOFEDCVLTATHLANTMDTRYIVL 1500
Qy 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAE 1560
Db 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLVGGGGGMLSRKGIIPBEYVLTRLAEDPAE 1560
Qy 1561 PRYTRRRRARFVSKKGCNCVAHKNIREQGRFLQDVFTTVLDLMPHTLLFTMSFLCSW 1620
Db 1561 PRYTRRRRARFVSKKGCNCVAHKNIREQGRFLQDVFTTVLDLMPHTLLFTMSFLCSW 1620
Qy 1621 LLFAMVWMLIAFAAGDLYAWEKGITDLAPRPVDPVYAGRGVPCVTSIHSSSAFLF 1680
Db 1621 LLFAMVWMLIAFAAGDLYAWEKGITDLAPRPVDPVYAGRGVPCVTSIHSSSAFLF 1680
Qy 1681 SIEVOVTIGFGRNATECEPLAILILIVONIVGIMINAMLGCFMKTAAQHRAEETLIF 1740
Db 1681 SIEVOVTIGFGRNATECEPLAILILIVONIVGIMINAMLGCFMKTAAQHRAEETLIF 1740
Qy 1741 SKPAVITLRGRCLCPMLRVGDLRKSMIISATIMQVVRKTSPEGEVVPVPLHOVDIPMENG 1800
Db 1741 SKPAVITLRGRCLCPMLRVGDLRKSMIISATIMQVVRKTSPEGEVVPVPLHOVDIPMENG 1800
Qy 1801 VGSNGIFLVAPLIIYHVIDSNSPLYDLAPSDLHHODEIIVILEGVETGITTQARTS 1860
Db 1801 VGSNGIFLVAPLIIYHVIDSNSPLYDLAPSDLHHODEIIVILEGVETGITTQARTS 1860
Qy 1861 YLADEIILMGGRFVPIVAEEDGRYSVDYSKFGNTIKVPTPLCTAQLDEDRSLDALTLAS 1920
Db 1861 YLADEIILMGGRFVPIVAEEDGRYSVDYSKFGNTIKVPTPLCTAQLDEDRSLDALTLAS 1920
Qy 1921 SRGPLRKRSVAVAKAPKFSISPDLS 1947
Db 1921 SRGPLRKRSVAVAKAPKFSISPDLS 1947

```

RESULT 4
 US-10-665-283-4
 ; Sequence 4, Application US/10665283
 ; Publication No. US2005063989A1
 ; GENERAL INFORMATION:
 ; APPLICANT: COMMISSARIAT A L'ENERGIE ATOMIQUE
 ; APPLICANT: DERAND, Renaud
 ; APPLICANT: GARCIA, Elisabeth

APPLICANT: PROST, Anne-Lise
APPLICANT: REVELLOU, Jean
APPLICANT: VIVANOU, Michel
TITLE OF INVENTION: ION CHANNEL HYBRID PROTEINS AS ELECTRICAL SENSORS AND USE THEREOF
FILE REFERENCE: P26305104
CURRENT APPLICATION NUMBER: US/10/665,283
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 1891
TYPE: PRT
ORGANISM: Homo sapiens
US-10-665-283-4

Query Match 96.9%; Score 9704; DB 5; Length 1891;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 20; Gaps 2;

1 MALGFCGADGSDPLMDNMTNTSNPDFTKCFONTLVWVPCFYLMACPFYFLYSRH 60
1 MALGFCGADGSDPLMDNMTNTSNPDFTKCFONTLVWVPCFYLMACPFYFLYSRH 60
61 DRGYIOMTPLNKTKTALGFLIMIVCWADLFYSFWERSRGIFLAFVFLVSPILLGITLLA 120
61 DRGYIOMTPLNKTKTALGFLIMIVCWADLFYSFWERSRGIFLAFVFLVSPILLGITLLA 120
121 TFLIOLERRKGVSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDFRDTFYVPS 180
121 TFLIOLERRKGVSSGIMLTFWLVALCALAILRSKIMTALKEDAQVDFRDTFYVPS 180
181 LLLIOLVSCFSDSPFSETHDNPCESSASFSLRTFWMITGILVRCYRQPLEGSD 240
181 LLLIOLVSCFSDSPFSETHDNPCESSASFSLRTFWMITGILVRCYRQPLEGSD 240
241 LMLINKEDTSEQVAVLVKMKKECAKTRKQPVKVVSSKDPAPKSSKVDANEVEAL 300
241 LMLINKEDTSEQVAVLVKMKKECAKTRKQPVKVVSSKDPAPKSSKVDANEVEAL 300
241 LMSLNKEDTSEQVAVLVKMKKECAKTRKQPVKVVSSKDPAPKSSKVDANEVEAL 300
301 IVKSPQKEMNPSLFKVLKTFGPFYFLMSFPFKAIHDLIMFSGPOLIKLIFVNDTRAPD 360
301 IVKSPQKEMNPSLFKVLKTFGPFYFLMSFPFKAIHDLIMFSGPOLIKLIFVNDTRAPD 360
361 MOGEFYVULFVTAACLOTIVLHOYFHI CFUSGMRKTRAVIGAVYRKALVITNSARKSTV 420
361 MOGEFYVULFVTAACLOTIVLHOYFHI CFUSGMRKTRAVIGAVYRKALVITNSARKSTV 420
421 GEIVNLMSVDAQRMDLATYINMIMISAPLOVITLALYLLMLNGPSYLAGVAVVLMVTPVN 480
421 GEIVNLMSVDAQRMDLATYINMIMISAPLOVITLALYLLMLNGPSYLAGVAVVLMVTPVN 480
481 AVNMAKTKTYOVAMHMSKDNRIKLMNEILINGIKVLYAMELAFKDKVLAIROBELKVLK 540
481 AVNMAKTKTYOVAMHMSKDNRIKLMNEILINGIKVLYAMELAFKDKVLAIROBELKVLK 540
541 KSAVLAVGFTWCTPFLVALCTFAVYVITDENNNIIDAQTAFLVALFNILRPPLNLP 600
541 KSAVLAVGFTWCTPFLVALCTFAVYVITDENNNIIDAQTAFLVALFNILRPPLNLP 600
601 MVISSIVQASVSLKRLRIFLSHEELPDSIERRPVKOGGGSITVENATFTMARSDPT 660
601 MVISSIVQASVSLKRLRIFLSHEELPDSIERRPVKOGGGSITVENATFTMARSDPT 660
661 LMGITTSIPREGALVAVVGQVCGKSSLSALLAEMDREGEHVALKGSVAVYPOQAWI QND 720
661 LMGITTSIPREGALVAVVGQVCGKSSLSALLAEMDREGEHVALKGSVAVYPOQAWI QND 720
721 SLRENIIFGCOLBEPPYRSYVIAOCALPDLIELPSGRTEIGEGVNLSSGQKORVSLAR 780
721 SLRENIIFGCOLBEPPYRSYVIAOCALPDLIELPSGRTEIGEGVNLSSGQKORVSLAR 780
781 AVYSNADIYLFDEPLASVDAHVGHIFENYIGPKMLKNKTRILVTHSMSTLPQVDVIIV 840
781 AVYSNADIYLFDEPLASVDAHVGHIFENYIGPKMLKNKTRILVTHSMSTLPQVDVIIV 840

781 AVYSNADIYLFDEPLASVDAHVGHIFENYIGPKMLKNKTRILVTHSMSTLPQVDVIIV 840
841 MSGGKIEMSGSYOELLARDGAPAFELRTYASTEBODAEENGVTGVSQPGKEAOMENGM 900
841 MSGGKIEMSGSYOELLARDGAPAFELRTYASTEBODAEENGVTGVSQPGKEAOMENGM 900
901 LVYDSAGKOLOROLSSSSSGDISRRHNSTAELOKAEAKKEEYWKMEADKAOTGOVKL 960
901 LVYDSAGKOLOROLSSSSSGDISRRHNSTAELOKAEAKKEEYWKMEADKAOTGOVKL 960
961 SVYWDYMKAIQLFISFLSIFLFCMCHVSALASNYMLSMTDDPIVNGTOEHTKVLRSYVG 1020
961 SVYWDYMKAIQLFISFLSIFLFCMCHVSALASNYMLSMTDDPIVNGTOEHTKVLRSYVG 1020
1021 ALGISOGIAVFGSMANVIGGILASRCLAYDLHSILRSFMSFPERIPSGULVNFSEKEL 1080
1021 ALGISOGIAVFGSMANVIGGILASRCLAYDLHSILRSFMSFPERIPSGULVNFSEKEL 1080
1081 DTVDSMIPEVIKMPGSLFNVIGACIVITLATTPIAIIIPPLGLIYFFVORFYVASSROL 1140
1081 DTVDSMIPEVIKMPGSLFNVIGACIVITLATTPIAIIIPPLGLIYFFVORFYVASSROL 1140
1141 KRLESVSRSPVYSHFNETLLGVSYIRAFEEQERFIHQSDIKVDENQKAYYPSIVANRWLA 1200
1141 KRLESVSRSPVYSHFNETLLGVSYIRAFEEQERFIHQSDIKVDENQKAYYPSIVANRWLA 1200
1201 VRLFCVNCIVLPAALFAVYSRHSLSAGLVGVSYSLOVTTYINMLVRMSSEMETNIVA 1260
1201 VRLFCVNCIVLPAALFAVYSRHSLSAGLVGVSYSLOVTTYINMLVRMSSEMETNIVA 1260
1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVFRNYCLRYREDLDFVLRHNVITNGG 1320
1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVFRNYCLRYREDLDFVLRHNVITNGG 1320
1321 EKVGIIVRTAGKSSLTGLFIRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
1321 EKVGIIVRTAGKSSLTGLFIRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
1381 SGSLRNMLDPPSOYSDEEVTSLLELAHLKDFVSALPDKLDHECAEGGENLSVQROVLVCL 1440
1381 SGSLRNMLDPPSOYSDEEVTSLLELAHLKDFVSALPDKLDHECAEGGENLSVQROVLVCL 1440
1441 ARALLRKTILVLDENAAVLDLTDLLOSTITQEPDCTVLTAHMLNTIMYTRIVL 1500
1441 ARALLRKTILVLDENAAVLDLTDLLOSTITQEPDCTVLTAHMLNTIMYTRIVL 1500
1501 DKGEIOEYGAUSDLOOQGLFYSMADAGLVGGGGGMLSRKGIIPREYVLTFLAEADPAE 1560
1501 DKGEIOEYGAUSDLOOQGLFYSMADAGLVGGGGGMLSRKGIIPREYVLTFLAEADPAE 1560
1561 PRYTRERRARFVSKKNCNVAAHKNIREQGRFLODVEFTVLVDLKWPHLLIIFMWSFLCSW 1620
1561 PRYTRERRARFVSKKNCNVAAHKNIREQGRFLODVEFTVLVDLKWPHLLIIFMWSFLCSW 1620
1621 LLEFAMWMLIAFAHGLYAMKGIYDLAPYDVPDYAGEGNNVPCVTSIHSPSSAFLF 1680
1621 LLEFAMWMLIAFAHGLYAMKGIYDLAPYDVPDYAGEGNNVPCVTSIHSPSSAFLF 1680
1681 SIEVQVITIGFGGMYTEECPLAILIVONIYIGMINAIMAGCI FMKTAOAHRAEFLIF 1740
1681 SIEVQVITIGFGGMYTEECPLAILIVONIYIGMINAIMAGCI FMKTAOAHRAEFLIF 1740
1741 SKIAVITLRRHRLCFMLRVGDLKSMISATTIMQVVRKTTSPGEVVPPLHVDIPIWENG 1800
1741 SKIAVITLRRHRLCFMLRVGDLKSMISATTIMQVVRKTTSPGEVVPPLHVDIPIWENG 1800
1801 VGGNGIFVAPLIIIVIVINSPLVDLAPSDLHHDDLEIIVLLEGVETTGITTOARTS 1860
1801 VGGNGIFVAPLIIIVIVINSPLVDLAPSDLHHDDLEIIVLLEGVETTGITTOARTS 1860
1861 YLADEILMGRFVPIVAEBDGRYSVDYSKRGNTIKVPTPLCTARQLEDEDS 1911
1861 YLADEILMGRFVPIVAEBDGRYSVDYSKRGNTIKVPTPLCTARQLEDEDS 1911
1911 YLADEILMGRFVPIVAEBDGRYSVDYSKRGNTIKVPTPLCTARQLEDEDS 1911
1911 YLADEILMGRFVPIVAEBDGRYSVDYSKRGNTIKVPTPLCTARQLEDEDS 1911

RESULT 5
US-09-939-853A-86
; Sequence 86, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939,853A
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 1531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-86

Query Match 78.5%; Score 7860; DB 3; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCASDSDPLMDNMVNTNTSNPDFTKCFQNTVLVWVPCFYLMACFPFYLYLSRH 60
DB 1 MALRGFCASDSDPLMDNMVNTNTSNPDFTKCFQNTVLVWVPCFYLMACFPFYLYLSRH 60
QY 61 DRGIYQNTPLNKTATLGFLLIMVCMADLFYSFWERSGCIPLAPVFLVSPLLGTTLLA 120
DB 61 DRGIYQNTPLNKTATLGFLLIMVCMADLFYSFWERSGCIPLAPVFLVSPLLGTTLLA 120
QY 121 TELLOLEBRKGVOSGIMLTFMVLVLCALATIRSKIMTALKEPDVNLFPDITFYVVS 180
DB 121 TELLOLEBRKGVOSGIMLTFMVLVLCALATIRSKIMTALKEPDVNLFPDITFYVVS 180
QY 121 TFLQLEBRKGVOSGIMLTFMVLVLCALATIRSKIMTALKEPDVNLFPDITFYVVS 180
DB 121 TFLQLEBRKGVOSGIMLTFMVLVLCALATIRSKIMTALKEPDVNLFPDITFYVVS 180
QY 181 LLLIQVLVSCFSDSPLESETIHDPNCPRESSASFLSRTTFMWITGLIVRGYRPLESSD 240
DB 181 LLLIQVLVSCFSDSPLESETIHDPNCPRESSASFLSRTTFMWITGLIVRGYRPLESSD 240
QY 241 LMSLNKEDTSQVVPVLVKNMKCECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVEAL 300
DB 241 LMSLNKEDTSQVVPVLVKNMKCECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVEAL 300
QY 301 IVKSPCKEMNPSLPVLYKTGPRVFLMSFPFKAHIDLMFSGPOLKLLIKFVNDTKAPD 360
DB 301 IVKSPCKEMNPSLPVLYKTGPRVFLMSFPFKAHIDLMFSGPOLKLLIKFVNDTKAPD 360
QY 361 WQGFYFYLLEFVTAQTLVLAHQYFHIQFVSGMRKTAIVIGAVYKALVITNSAKSSSTV 420
DB 361 WQGFYFYLLEFVTAQTLVLAHQYFHIQFVSGMRKTAIVIGAVYKALVITNSAKSSSTV 420
QY 421 GEIYVNLMSVDAQRFMDLATYINMISAPLOVILALYLLMLNLGSPVLGAVAMVIMEVN 480
DB 421 GEIYVNLMSVDAQRFMDLATYINMISAPLOVILALYLLMLNLGSPVLGAVAMVIMEVN 480
QY 481 AVMAKTKTYOVAAHKSKDNRIKLMNELNGIKYLKIYAMELAFOKYLAIROEELKYLK 540
DB 481 AVMAKTKTYOVAAHKSKDNRIKLMNELNGIKYLKIYAMELAFOKYLAIROEELKYLK 540
QY 541 KSAVLSAVGTFVWCTPFLVALCTPAVYVTTIDENNIILAOQAFVSLALFNILRPPLNTLP 600
DB 541 KSAVLSAVGTFVWCTPFLVALCTPAVYVTTIDENNIILAOQAFVSLALFNILRPPLNTLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHBELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660

DB 601 MVISSIVQASVSLKRLRIFLSHBELEPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660
QY 661 LMGITFSEIPREGALVAVVGVQCCGKSSLSLALAMEDKYEGVAIKGSVAVYPQOAWIOND 720
DB 661 LMGITFSEIPREGALVAVVGVQCCGKSSLSLALAMEDKYEGVAIKGSVAVYPQOAWIOND 720
QY 721 SLRENILFGCQLEBPYRSVYIQAACALLPDEILPDSGRTEIGEKVNLGGQKORVSLAR 780
DB 721 SLRENILFGCQLEBPYRSVYIQAACALLPDEILPDSGRTEIGEKVNLGGQKORVSLAR 780
QY 781 AVYNNADITYLEPDDLAVDAVGHIFENYIGPQMKLNKTRILLVTHSMSTLPQVDVYIV 840
DB 781 AVYNNADITYLEPDDLAVDAVGHIFENYIGPQMKLNKTRILLVTHSMSTLPQVDVYIV 840
QY 841 MSGGKISEMGSYQELALRDGAFAFLRTYASTEOEODAEENGVTGVSQPGKEALOMENGM 900
DB 841 MSGGKISEMGSYQELALRDGAFAFLRTYASTEOEODAEENGVTGVSQPGKEALOMENGM 900
QY 901 LVYDSAGKQLOROLSSSSSYSGDISRHHNSTAELQKAEXKEETWKLMEADKAOTGOVKL 960
DB 901 LVYDSAGKQLOROLSSSSSYSGDISRHHNSTAELQKAEXKEETWKLMEADKAOTGOVKL 960
QY 961 SVYWDYMKATGLFISFLSIFLPMCNHVSALASNYWLSLMTDDPIVNGTQETHKRLSYVG 1020
DB 961 SVYWDYMKATGLFISFLSIFLPMCNHVSALASNYWLSLMTDDPIVNGTQETHKRLSYVG 1020
QY 1021 ALGISOGIAVFGYMAVSIIGIILASRCIHLVDLHLSIRSPMSFEPTPSGNLVNRFSEL 1080
DB 1021 ALGISOGIAVFGYMAVSIIGIILASRCIHLVDLHLSIRSPMSFEPTPSGNLVNRFSEL 1080
QY 1081 DTVDMSIPEVIMKMGSLFNVIAGACIVILLATPIAIIIPPLGIYEFVQGFYVASSRQL 1140
DB 1081 DTVDMSIPEVIMKMGSLFNVIAGACIVILLATPIAIIIPPLGIYEFVQGFYVASSRQL 1140
QY 1141 KRLSVSRSPYVSHFNETLGVSYIRAFEEQERFIHOSDLKVDENOKAYYSIVANRWLA 1200
DB 1141 KRLSVSRSPYVSHFNETLGVSYIRAFEEQERFIHOSDLKVDENOKAYYSIVANRWLA 1200
QY 1201 VRLSCVNGCIVLPALRAVVISRHSLSAGLVLSYSIQVTTYINMLVRRMSSEMETNIVA 1260
DB 1201 VRLSCVNGCIVLPALRAVVISRHSLSAGLVLSYSIQVTTYINMLVRRMSSEMETNIVA 1260
QY 1261 VERLKEVSETEKEAPWQIETAPPSWPQVGRVFRNYCLATYRDDLPVLAHINVTINGG 1320
DB 1261 VERLKEVSETEKEAPWQIETAPPSWPQVGRVFRNYCLATYRDDLPVLAHINVTINGG 1320
QY 1321 EKVGI VGRGTGAKSSLTGLFRINESABGEIIDGINIAKIGLHDLRKTIIIPQDPVLF 1380
DB 1321 EKVGI VGRGTGAKSSLTGLFRINESABGEIIDGINIAKIGLHDLRKTIIIPQDPVLF 1380
QY 1381 SGLSRMLNDPFSQVSDSEEWMTSLBLAHKDPVSAIPKLDHECAEGGENLSVGOROLVCL 1440
DB 1381 SGLSRMLNDPFSQVSDSEEWMTSLBLAHKDPVSAIPKLDHECAEGGENLSVGOROLVCL 1440
QY 1441 ARAILRKTKILVDEATAVADLETTDDLIQSTIRTOFEDCTVLTIAHRLNTIMDTRVIVL 1500
DB 1441 ARAILRKTKILVDEATAVADLETTDDLIQSTIRTOFEDCTVLTIAHRLNTIMDTRVIVL 1500
QY 1501 DKGEI QEYGA PSDLLQORGLFYMAKDAGLV 1531
DB 1501 DKGEI QEYGA PSDLLQORGLFYMAKDAGLV 1531

RESULT 6
US-10-667-891-6
; Sequence 6, Application US/10667891
; Publication No. US20040171024A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, CHARLES W.
; APPLICANT: BRET, PAUL T.
; APPLICANT: HOLM, INGE
; APPLICANT: GRAILLES, MARINE

APPLICANT: RZHETSKY, ANDREY
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND
FILE REFERENCE: 03495.0294-00000
CURRENT APPLICATION NUMBER: US/10/667,891
CURRENT FILING DATE: 2003-09-23
PRIOR APPLICATION NUMBER: 60/413,469
PRIOR FILING DATE: 2002-09-26
NUMBER OF SEQ ID NOS: 76
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO: 6
LENGTH: 1531
TYPE: PRT
ORGANISM: Homo sapiens
US-10-667-891-6

Query Match 78.5%; Score 7860; DB 4; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRPGSADGSDPLMDNVTMTNSNDPTKQONTLVVWPCFYLMACFPFYLYSRH 60
DB 1 MALRPGSADGSDPLMDNVTMTNSNDPTKQONTLVVWPCFYLMACFPFYLYSRH 60
QY 61 DRGYIOMTPLNKTALGFLIMIVCMADLFYSPWERSRGIFLAIVFLVSPFLIGITLLA 120
DB 61 DRGYIOMTPLNKTALGFLIMIVCMADLFYSPWERSRGIFLAIVFLVSPFLIGITLLA 120
QY 121 TFLIQLERRKGVSGGIMLTFWVALVCAALILRSKIMTALKEDAQVDFRDIIFYYFS 180
DB 121 TFLIQLERRKGVSGGIMLTFWVALVCAALILRSKIMTALKEDAQVDFRDIIFYYFS 180
QY 181 LLLIQLVLSGCSDDSPLESETIHDNPPCESSASFLSRITWMTTGLIVGRPLSGSD 240
DB 181 LLLIQLVLSGCSDDSPLESETIHDNPPCESSASFLSRITWMTTGLIVGRPLSGSD 240
QY 241 LMSLNKEDTSEQVVPVVKWKECAKTRKOPVAVVSSKDPAPKSSKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVVKWKECAKTRKOPVAVVSSKDPAPKSSKVDANEVEAL 300
QY 301 IVKSPQKMNPSLFKVLKTFPGYFLMSFFPKAIDHLMFSGPQILKLLIKFVNDTRAPD 360
DB 301 IVKSPQKMNPSLFKVLKTFPGYFLMSFFPKAIDHLMFSGPQILKLLIKFVNDTRAPD 360
QY 361 WQGFYVVLFTYACLOTVLHOFYHICFVSGMARIKAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYVVLFTYACLOTVLHOFYHICFVSGMARIKAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIYNLMSVDAOREMDLATYINMWSAPLOYITLALYLLMLNIGPSVLAGVAVVWMLMPVN 480
DB 421 GEIYNLMSVDAOREMDLATYINMWSAPLOYITLALYLLMLNIGPSVLAGVAVVWMLMPVN 480
QY 481 AVAMKTKTYOVAMHMSKDNRIKLMNEILINGIKVLUKYLAMELAFKDYLAIROBELKVLK 540
DB 481 AVAMKTKTYOVAMHMSKDNRIKLMNEILINGIKVLUKYLAMELAFKDYLAIROBELKVLK 540
QY 541 KSAVLAVGFTWVCTPFLVALCTFAVYVITDENNNIDAQTAFTVSLFNLREPLNLP 600
DB 541 KSAVLAVGFTWVCTPFLVALCTFAVYVITDENNNIDAQTAFTVSLFNLREPLNLP 600
QY 601 MVTSSIVQASVSLKRLIFLSHELEPDSIERRPVKGSGGNSITVANAFTPARSDPT 660
DB 601 MVTSSIVQASVSLKRLIFLSHELEPDSIERRPVKGSGGNSITVANAFTPARSDPT 660
QY 661 LINGITFIPGALVAVVGOVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAWIOND 720
DB 661 LINGITFIPGALVAVVGOVCGKSSLLSALLAEMDKVEGHVAIKGSVAVYVPOQAWIOND 720
QY 721 SLRENIIFGQLEPPYRSYIQAACALLPDLIELPSGRTIEGEGVNLSSGQKQVSLAR 780
DB 721 SLRENIIFGQLEPPYRSYIQAACALLPDLIELPSGRTIEGEGVNLSSGQKQVSLAR 780
QY 781 AVYSNADIYFDPLSAVDHVGHIPEVNIIGPKGMKNKTRILVTHSMGYLPQVDYIIV 840

DB 781 AVYSNADIYFDPLSAVDHVGHIPEVNIIGPKGMKNKTRILVTHSMGYLPQVDYIIV 840
QY 841 MSGGKISEMSYQELARDAFAFLRTYASTEOBDAENGVTGVSQPKKAKOMENG 900
DB 841 MSGGKISEMSYQELARDAFAFLRTYASTEOBDAENGVTGVSQPKKAKOMENG 900
QY 901 LYTDSAGKOLQORLSSSSYSYSDISRHHNSTAELOKAEKKEFTWKLMEADKQOTGVKL 960
DB 901 LYTDSAGKOLQORLSSSSYSYSDISRHHNSTAELOKAEKKEFTWKLMEADKQOTGVKL 960
QY 961 SYWYDVKAIGLFISFLIFLFCNHNVSALASNYMLMTDPIVNGTOEHTKRLSVYG 1020
DB 961 SYWYDVKAIGLFISFLIFLFCNHNVSALASNYMLMTDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISQIAVFGYSMAVSIIGIILASRCLHYDLHSIIRSPMSFERTPSGNLVNRSKEL 1080
DB 1021 ALGISQIAVFGYSMAVSIIGIILASRCLHYDLHSIIRSPMSFERTPSGNLVNRSKEL 1080
QY 1081 DTVDSMIPEYIKMFMGSLFNVIACIVILATPLAIIIPPLGIYFVQRPVASSROL 1140
DB 1081 DTVDSMIPEYIKMFMGSLFNVIACIVILATPLAIIIPPLGIYFVQRPVASSROL 1140
QY 1141 KRLESYRSRSPVSHFNETLIGSVIRAFEEQERFIHQSDIKVDENQKAYPSIVANMILA 1200
DB 1141 KRLESYRSRSPVSHFNETLIGSVIRAFEEQERFIHQSDIKVDENQKAYPSIVANMILA 1200
QY 1201 VRLCEVNCIVLPALPAVIRSHLSAGLVGSVSVLOVTTYANLVNRSSEMETIVA 1260
DB 1201 VRLCEVNCIVLPALPAVIRSHLSAGLVGSVSVLOVTTYANLVNRSSEMETIVA 1260
QY 1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVEFRNYCIRYEDLDPUVRHNVITNGG 1320
DB 1261 VERLKEYSETEKAPMOIOETAPSSWPQVGRVEFRNYCIRYEDLDPUVRHNVITNGG 1320
QY 1321 EKVIGVGTAGKSSLLTGLFRINESAGEIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
DB 1321 EKVIGVGTAGKSSLLTGLFRINESAGEIIDGINIAKIGLHDLRFKTTIIPQDPVLF 1380
QY 1381 SSSLRNMLDPSQYSDBEVWTSLELAHLKDFVSALPDKLDHECAGEGENISVQORQVLC 1440
DB 1381 SSSLRNMLDPSQYSDBEVWTSLELAHLKDFVSALPDKLDHECAGEGENISVQORQVLC 1440
QY 1441 ARALLKTKTLVDEATAADLETDDLIQSTITQEBDCVLTARLNTIMYTRIVL 1500
DB 1441 ARALLKTKTLVDEATAADLETDDLIQSTITQEBDCVLTARLNTIMYTRIVL 1500
QY 1501 DKGEIOYGAPSDILOORGIFYSMADAGLV 1531
DB 1501 DKGEIOYGAPSDILOORGIFYSMADAGLV 1531

RESULT 7
US-10-807-466-6
Sequence 6, Application US/10807466
Publication No. US20040244066A1
GENERAL INFORMATION:
APPLICANT: ROTH, CHARLES W.
APPLICANT: BREY, PAUL T.
APPLICANT: HOLM, INGE
APPLICANT: GRALLES, MARINE
APPLICANT: RZHETSKY, ANDREY
TITLE OF INVENTION: MULTIDRUG RESISTANCE PROTEINS IN DROSOPHILA AND
FILE REFERENCE: 03495.0294-01000
CURRENT APPLICATION NUMBER: US/10/807,466
CURRENT FILING DATE: 2004-03-24
PRIOR APPLICATION NUMBER: 10/667,891
PRIOR FILING DATE: 2003-09-23
PRIOR APPLICATION NUMBER: 60/413,469
PRIOR FILING DATE: 2002-09-26
NUMBER OF SEQ ID NOS: 139
SOFTWARE: PatentIn Ver. 3.2

SEQ ID NO 6
; LENGTH: 1531
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-807-466-6

Query Match 78.5%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MALRGFSGADSDPLMDNMVNTMTNSPDFTKCFQNTVLMVPCFYLMACFPFFLYLSRH 60
DB 1 MALRGFSGADSDPLMDNMVNTMTNSPDFTKCFQNTVLMVPCFYLMACFPFFLYLSRH 60
QY 61 DRGYIOMTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLAAPFLVSPILLGITTLA 120
DB 61 DRGYIOMTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLAAPFLVSPILLGITTLA 120
QY 121 TPLIQLEERRKGVSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYS 180
DB 121 TPLIQLEERRKGVSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDLFRDITFYVYS 180
QY 181 LLLIQLVISCFSDSPLESETIHDPNCPESASFLSRITFMWITGLIVRGYROPLESD 240
DB 181 LLLIQLVISCFSDSPLESETIHDPNCPESASFLSRITFMWITGLIVRGYROPLESD 240
QY 241 LMSLINKEDTSQVVPVLVKNMKKECAKTRKQPVKVYSSKDPAPCKESSKYDANEVEAL 300
DB 241 LMSLINKEDTSQVVPVLVKNMKKECAKTRKQPVKVYSSKDPAPCKESSKYDANEVEAL 300
QY 301 IVKSPQKEMNPSLFVLYKTGPTPLMSPPFKAIHDLMMFSGPQILKLIKRVNDTKAPD 360
DB 301 IVKSPQKEMNPSLFVLYKTGPTPLMSPPFKAIHDLMMFSGPQILKLIKRVNDTKAPD 360
QY 361 MGGFYTYTLTFTVCTLOTLVLMHGYHICFVSGMRKKTAVIGAVYKALVITNSAKSSTV 420
DB 361 MGGFYTYTLTFTVCTLOTLVLMHGYHICFVSGMRKKTAVIGAVYKALVITNSAKSSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINMISAPLOVILALYLMNLGSPVLAVAMVLMVAVN 480
DB 421 GEIVNLSVDAQRFMDLATYINMISAPLOVILALYLMNLGSPVLAVAMVLMVAVN 480
QY 481 AVMAKTKTYVVAHMKSKDNRIKLMNEILNGIKVILKYAMEIAFKQVLAIROBELKYLK 540
DB 481 AVMAKTKTYVVAHMKSKDNRIKLMNEILNGIKVILKYAMEIAFKQVLAIROBELKYLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTFAVVTIDENNILDAQTAFFVSLAFNILLRPLNLP 600
DB 541 KSAVLSAVGTFTWCTPFLVALCTFAVVTIDENNILDAQTAFFVSLAFNILLRPLNLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPVKDGGTNSITVRNATFTWASDPPT 660
QY 661 LMGITFSIPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAKGSVAAYVPOQAMIOND 720
DB 661 LMGITFSIPEGALVAVVGQVCGKSSLLSALLAEMDKVEGHVAKGSVAAYVPOQAMIOND 720
QY 721 SLRENIILFGCOLLEERYRSVIOACALLPDLLEILPSGDPTEIGEKVNLSSGQKORVSLAR 780
DB 721 SLRENIILFGCOLLEERYRSVIOACALLPDLLEILPSGDPTEIGEKVNLSSGQKORVSLAR 780
QY 781 AVYENADIYLFDDPLSAVDAAVGHKIFENVIGPKMKLNKRIILVTHSMSTYLPQVDVIV 840
DB 781 AVYENADIYLFDDPLSAVDAAVGHKIFENVIGPKMKLNKRIILVTHSMSTYLPQVDVIV 840
QY 841 MSGKISMSGSYOEILLARDGAFAEFLRTYASTEOBDAEENGVTGVSQPGKEAKOMENG 900
DB 841 MSGKISMSGSYOEILLARDGAFAEFLRTYASTEOBDAEENGVTGVSQPGKEAKOMENG 900
QY 901 LVTSAGAKOLROLLSSSSSSSGDISSRHNSITAELOKAKAKKEETKMLEADKAKQGVYKL 960
DB 901 LVTSAGAKOLROLLSSSSSSSGDISSRHNSITAELOKAKAKKEETKMLEADKAKQGVYKL 960
```

```
QY 961 SVYWDYKKAIGLFISFLSIFLFMCNHYSAALASNYMLSLMTDDPIVNGTOEHTKYRLSVYG 1020
DB 961 SVYWDYKKAIGLFISFLSIFLFMCNHYSAALASNYMLSLMTDDPIVNGTOEHTKYRLSVYG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLHYDILHSILRSPMSFFERTPSGNLVNRFSEKL 1080
DB 1021 ALGISOGIAVFGYSMAVSIIGIILASRCLHYDILHSILRSPMSFFERTPSGNLVNRFSEKL 1080
QY 1081 DTVDMSIPEVIMKMGSLFNVIGACIVILATPIAIIIPPLGLIYFVQRFYVASSROL 1140
DB 1081 DTVDMSIPEVIMKMGSLFNVIGACIVILATPIAIIIPPLGLIYFVQRFYVASSROL 1140
QY 1141 KRLSEVSRSPPYSHFNLTLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLSEVSRSPPYSHFNLTLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLCEVGNCTIVLPALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETNIVA 1260
DB 1201 VRLCEVGNCTIVLPALFAVISRHSLSAGLVGLSVSYSLQVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMOIOETAPPSMPOVGRVFRNYCLRYREDLPVLRHINVTINGG 1320
DB 1261 VERLKEYSETEKAPMOIOETAPPSMPOVGRVFRNYCLRYREDLPVLRHINVTINGG 1320
QY 1321 EKVGIVKRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIVKRTGAGKSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SSGSRMNLDPFSQYSDEBWTSLBLAHLKDFVSAHPKLDHECAGEBENISVGOROLVCL 1440
DB 1381 SSGSRMNLDPFSQYSDEBWTSLBLAHLKDFVSAHPKLDHECAGEBENISVGOROLVCL 1440
QY 1441 ABALIRKTKIIVLDEAPAAVDLEETDILQSTIRFQFEDCTVLTAAHRLNTIMDTRYIVL 1500
DB 1441 ABALIRKTKIIVLDEAPAAVDLEETDILQSTIRFQFEDCTVLTAAHRLNTIMDTRYIVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531
DB 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531

RESULT 8
US-10-484-577-678
; Sequence 678, Application US/10484577
; Publication No. US20050032724A1
; GENERAL INFORMATION:
; APPLICANT: EPIDAUROS Biotechnologie Aktiengesellschaft
; TITLE OF INVENTION: Means and methods for improved treatment of cancer based on UGT1A1
; FILE REFERENCE: P2285PCT-1
; CURRENT APPLICATION NUMBER: US/10/484,577
; CURRENT FILING DATE: 2004-01-22
; PRIOR APPLICATION NUMBER: PCT/EP 02/08220
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: EP 01 11 7608.8
; PRIOR FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: EP 02011710.7
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 683
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 678
; LENGTH: 1531
; TYPE: PRN
; ORGANISM: Homo sapiens
US-10-484-577-678

Query Match 78.5%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MALRGFSGADSDPLMDNMVNTMTNSPDFTKCFQNTVLMVPCFYLMACFPFFLYLSRH 60
DB 1 MALRGFSGADSDPLMDNMVNTMTNSPDFTKCFQNTVLMVPCFYLMACFPFFLYLSRH 60
```

QY 61 DRGIOMTPANKTKTALGFLMIWCMADLFYSFWMERSRGIFLAPVFLVSPILLGITTLLA 120
DB 61 DRGIOMTPANKTKTALGFLMIWCMADLFYSFWMERSRGIFLAPVFLVSPILLGITTLLA 120
QY 121 TFLIOLERRKGVSSGIMLTFMLVVALVCALALRSKIMTALKEDAQVDLFRDITFFYYFS 180
DB 121 TFLIOLERRKGVSSGIMLTFMLVVALVCALALRSKIMTALKEDAQVDLFRDITFFYYFS 180
QY 181 LLLIOLVLSGSDSPLESETIHDPNCPRESSASLSTITFWMTGLIVRGYRQPLESSD 240
DB 181 LLLIOLVLSGSDSPLESETIHDPNCPRESSASLSTITFWMTGLIVRGYRQPLESSD 240
QY 241 LMSLNKEDTSBOQVPLVVKMKKECAKTRKQPVKVVYSSKDPAPKRESSKVDANAEVEAL 300
DB 241 LMSLNKEDTSBOQVPLVVKMKKECAKTRKQPVKVVYSSKDPAPKRESSKVDANAEVEAL 300
QY 301 IVKSPKEMWNSLFKVLKTFPGPYFLMSFFKALHDLMPFSGPOLKLLIFVNDTKAPD 360
DB 301 IVKSPKEMWNSLFKVLKTFPGPYFLMSFFKALHDLMPFSGPOLKLLIFVNDTKAPD 360
QY 361 WQGFYTVLFTVACLOTLVHQYFHCIFVSGMKIKTAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYTVLFTVACLOTLVHQYFHCIFVSGMKIKTAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDAQRFMDLATYINMIMSAPLOVIALYILMLNLSGSLVAGAVMTLMPVN 480
DB 421 GEIVNLMSVDAQRFMDLATYINMIMSAPLOVIALYILMLNLSGSLVAGAVMTLMPVN 480
QY 481 AYMAKTKTYOVAMHMSKDNRIKLMNEILNGIKVLKYAMELAFKDKVLAIROBELKVLK 540
DB 481 AYMAKTKTYOVAMHMSKDNRIKLMNEILNGIKVLKYAMELAFKDKVLAIROBELKVLK 540
QY 541 KSAVLSAVGTFTWCTPPLVALCTPAAVYITDENNILDQAFAVSLAFNILRPLNLP 600
DB 541 KSAVLSAVGTFTWCTPPLVALCTPAAVYITDENNILDQAFAVSLAFNILRPLNLP 600
QY 601 MVLSIYOASVSLKRLIFLSHEELPDSIERRPVKOGGNGSTTVNATPTMARSPPT 660
DB 601 MVLSIYOASVSLKRLIFLSHEELPDSIERRPVKOGGNGSTTVNATPTMARSPPT 660
QY 661 LMGTFESIPRGALVAVVGOCGKSSLSALLAEMDKVGHVAKGSVAVYPOQAMQND 720
DB 661 LMGTFESIPRGALVAVVGOCGKSSLSALLAEMDKVGHVAKGSVAVYPOQAMQND 720
QY 721 SLRENILFGCOLPEPYRSVIOACALLPDEILPSGRTIEGKGNVLSGGQKQVSLAR 780
DB 721 SLRENILFGCOLPEPYRSVIOACALLPDEILPSGRTIEGKGNVLSGGQKQVSLAR 780
QY 781 AVYSNADIYLPDDPLSAVDHVGKIFENYIGPKGMKNTKRIIVTHSMSTLPQVDYIV 840
DB 781 AVYSNADIYLPDDPLSAVDHVGKIFENYIGPKGMKNTKRIIVTHSMSTLPQVDYIV 840
QY 841 MSGGKISEMGSYOELARDAFAFLRTYASTEOQDAEENGVTGSGPKGAQOMENGM 900
DB 841 MSGGKISEMGSYOELARDAFAFLRTYASTEOQDAEENGVTGSGPKGAQOMENGM 900
QY 901 LVTDSAGKOLORQLSSSSSYSGDISRRHNSTALIOKAKEKETWKLMEADKAQTQVXL 960
DB 901 LVTDSAGKOLORQLSSSSSYSGDISRRHNSTALIOKAKEKETWKLMEADKAQTQVXL 960
QY 961 SYVMDVMKATIGLFSFSLIFLMCNHVSALASNYLSLMDPDPVNGTOEHTKRLSVYG 1020
DB 961 SYVMDVMKATIGLFSFSLIFLMCNHVSALASNYLSLMDPDPVNGTOEHTKRLSVYG 1020
QY 1021 ALGISOGIAVFGVMAVSIIGIILASRCILHVDLHSLIRSPMSFFERTPSGNLVNRFSEKL 1080
DB 1021 ALGISOGIAVFGVMAVSIIGIILASRCILHVDLHSLIRSPMSFFERTPSGNLVNRFSEKL 1080
QY 1081 DYTDSMTPEVYKMGSLFNVIKACIYIILATPIAIIIPPLGIYTFVQRFYVASSROL 1140
DB 1081 DYTDSMTPEVYKMGSLFNVIKACIYIILATPIAIIIPPLGIYTFVQRFYVASSROL 1140
QY 1141 KRLSEVSRSPYSHFNETLIGSVIRAFEEQEREIHOSDKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLSEVSRSPYSHFNETLIGSVIRAFEEQEREIHOSDKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLCEVNCIYLFALFAVISRHSLSAGVLVSYSLOVTTYLNMVLVRMSSEMETNIVA 1260
DB 1201 VRLCEVNCIYLFALFAVISRHSLSAGVLVSYSLOVTTYLNMVLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMOIOETAPPSWPOVGRVFRNYCLARREDLPVLRHINVTNGG 1320
DB 1261 VERLKEYSETEKAPMOIOETAPPSWPOVGRVFRNYCLARREDLPVLRHINVTNGG 1320
QY 1321 EKVIGVGTAGKSLTLGFRINESAGEIIGINIAKIGLHDLRFKTIIPQDPVLF 1380
DB 1321 EKVIGVGTAGKSLTLGFRINESAGEIIGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGLRNMIDPFSSQSDSEWVTSLELAHKDFVSALPDKLDBECAGEENLSVGROQLVCL 1440
DB 1381 SGLRNMIDPFSSQSDSEWVTSLELAHKDFVSALPDKLDBECAGEENLSVGROQLVCL 1440
QY 1441 ARALIRTKTILVDEATAVADLETDLLOSTIRQFEDCTVLTIAHRLNTIMDYTRIVL 1500
DB 1441 ARALIRTKTILVDEATAVADLETDLLOSTIRQFEDCTVLTIAHRLNTIMDYTRIVL 1500
QY 1501 DKGEIOEYGA PSDLQORGLFYSMAKXGALV 1531
DB 1501 DKGEIOEYGA PSDLQORGLFYSMAKXGALV 1531

RESULT 9
US-10-889-503-19
Sequence 19, Application US/10889503
Publication No. US20050063968A1
GENERAL INFORMATION:
APPLICANT: Fox Chase Cancer Center
APPLICANT: Krush, Gary D.
APPLICANT: Lee, Kun
APPLICANT: Belinsky, Martin G.
TITLE OF INVENTION: MRP-Related ABC Transporter Encoding
FILE REFERENCE: FCCC 98-02
CURRENT APPLICATION NUMBER: US/10/889,503
PRIOR FILING DATE: 2004-07-12
PRIOR APPLICATION NUMBER: US/09/647,140
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US99/06644
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 60/079,759
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/095,153
PRIOR FILING DATE: 1998-08-03
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 19
LENGTH: 1531
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-889-503-19
Query Match 78.5%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSADGSDPLMDMNTVNTSNPDFKCFQNTLVVWPCFYLMACPPFFLYLSRH 60
DB 1 MALRGFCSADGSDPLMDMNTVNTSNPDFKCFQNTLVVWPCFYLMACPPFFLYLSRH 60
QY 61 DRGIOMTPANKTKTALGFLMIWCMADLFYSFWMERSRGIFLAPVFLVSPILLGITTLLA 120
DB 61 DRGIOMTPANKTKTALGFLMIWCMADLFYSFWMERSRGIFLAPVFLVSPILLGITTLLA 120
QY 121 TFLIOLERRKGVSSGIMLTFMLVVALVCALALRSKIMTALKEDAQVDLFRDITFFYYFS 180
DB 121 TFLIOLERRKGVSSGIMLTFMLVVALVCALALRSKIMTALKEDAQVDLFRDITFFYYFS 180

```
|||||
Db 121 TFLQLERRKGVQSSGIMLTFMLVALVCALAILRSKIMTALKEDAOVDLFFRDIFFYVFS 180
Qy 181 LLLIQLVLSGSDSPLESETIHDNPNCPSSASFLSITFMWITGLIVRGROPLEGSD 240
Db 181 LLLIQLVLSGSDSPLESETIHDNPNCPSSASFLSITFMWITGLIVRGROPLEGSD 240
Qy 241 LMSLNKEDTSQOVVPLVKNMKKECAKTRKQPVKVYVSSKDPAPCKESSKYDANEVEAL 300
Db 241 LMSLNKEDTSQOVVPLVKNMKKECAKTRKQPVKVYVSSKDPAPCKESSKYDANEVEAL 300
Qy 301 IVKSPQKEMNPSLFKVLYKTGPFYFLMSFPFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360
Db 301 IVKSPQKEMNPSLFKVLYKTGPFYFLMSFPFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360
Qy 361 WQGFYTYLLFVTACLOTLVLHQYFHI CFVSGMR IKTAIVIGAVYRKALVITNSARKSSTV 420
Db 361 WQGFYTYLLFVTACLOTLVLHQYFHI CFVSGMR IKTAIVIGAVYRKALVITNSARKSSTV 420
Qy 421 GEIYVNLMSVDAQRFMDLATYINMIMSAPLOVILALYLLMLNGPSVLGAVAVMVLMBEVN 480
Db 421 GEIYVNLMSVDAQRFMDLATYINMIMSAPLOVILALYLLMLNGPSVLGAVAVMVLMBEVN 480
Qy 481 AVAMAMKTTYOVAAHMKSKONRIKLMNEILNGIKVILKIYAMBLAFKDKVLAIRQBELKVLYK 540
Db 481 AVAMAMKTTYOVAAHMKSKONRIKLMNEILNGIKVILKIYAMBLAFKDKVLAIRQBELKVLYK 540
Qy 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTDENNIILOAQAFAVSLAFNIIIRPPLNIIIP 600
Db 541 KSAVLSAVGTFTWCTPFLVALCTFAVYVTTDENNIILOAQAFAVSLAFNIIIRPPLNIIIP 600
Qy 601 MVISIIVQASVSLKRLRIFLSHEBELPDSIERRPVKDGGGNTSITVRNATFTWASDDEPT 660
Db 601 MVISIIVQASVSLKRLRIFLSHEBELPDSIERRPVKDGGGNTSITVRNATFTWASDDEPT 660
Qy 661 LMGITFSLPEBAGLVAVVQVQVCGKSSLLSALLAENDKVBGHVATKGSVAVYPOQAMIOND 720
Db 661 LMGITFSLPEBAGLVAVVQVQVCGKSSLLSALLAENDKVBGHVATKGSVAVYPOQAMIOND 720
Qy 721 SLRENIILGCOLLEBEYYRSVIOACALLPDLLETLPBGSDTEIGEKKVNLSSGGKQKQVSLAR 780
Db 721 SLRENIILGCOLLEBEYYRSVIOACALLPDLLETLPBGSDTEIGEKKVNLSSGGKQKQVSLAR 780
Qy 781 AVYSNADIIYLPDDPLSAVDAAHVGHKAI FENVIGPKMKNKRIILVTHSMSYLPQYDVIIV 840
Db 781 AVYSNADIIYLPDDPLSAVDAAHVGHKAI FENVIGPKMKNKRIILVTHSMSYLPQYDVIIV 840
Qy 841 MSGKISMGSYOELLARDGAFAEFLRTYASTEBOQDABENGVTGVSQPKAKOMENGM 900
Db 841 MSGKISMGSYOELLARDGAFAEFLRTYASTEBOQDABENGVTGVSQPKAKOMENGM 900
Qy 901 LVTSAGKOLORUSSSSSYSGDISRHNSHTAELOKAEKKKEBTMKLEADAOQGOYVL 960
Db 901 LVTSAGKOLORUSSSSSYSGDISRHNSHTAELOKAEKKKEBTMKLEADAOQGOYVL 960
Qy 961 SVYMDYMAKIGLIFSLIFLFLMCHVBSALSNYMLSMITDDPIVNGTOEHTKVALSYVG 1020
Db 961 SVYMDYMAKIGLIFSLIFLFLMCHVBSALSNYMLSMITDDPIVNGTOEHTKVALSYVG 1020
Qy 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLHLSILNSPMSFFERTPSGNILVNFPSKEL 1080
Db 1021 ALGISOGIAVFGYMAVSIIGIILASRCLHVDLHLSILNSPMSFFERTPSGNILVNFPSKEL 1080
Qy 1081 DTVSMITEVIMKMFNGSLFENVIGACIVILLATFPIAIIIPRLGIYFPVQRYVASSQOL 1140
Db 1081 DTVSMITEVIMKMFNGSLFENVIGACIVILLATFPIAIIIPRLGIYFPVQRYVASSQOL 1140
Qy 1141 KRLSEVSRSPVYSHNETLIGSVIRAFEBEORFIHQSDLKVDENOKAYPSIVANRMILA 1200
Db 1141 KRLSEVSRSPVYSHNETLIGSVIRAFEBEORFIHQSDLKVDENOKAYPSIVANRMILA 1200
Qy 1201 VRLBVCNCIVLFAALPAVISRHSISAGIVGLSVSYSLQVTTYLNMVLVMSSEMETNIVA 1260
Db 1201 VRLBVCNCIVLFAALPAVISRHSISAGIVGLSVSYSLQVTTYLNMVLVMSSEMETNIVA 1260
```

```
Db 1201 VRLBVCNCIVLFAALPAVISRHSISAGIVGLSVSYSLQVTTYLNMVLVMSSEMETNIVA 1260
Qy 1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVFRNRYCLRYREDDLVJLHINVTINGG 1320
Db 1261 VERLKEYSETEKEAPMOIOETAPSSWPQVGRVFRNRYCLRYREDDLVJLHINVTINGG 1320
Qy 1321 EKVGIVGRTGAKGSLTLGLFRINESAGEIIDIINIAKIGLHDLRKTTIIPQDPVLF 1380
Db 1321 EKVGIVGRTGAKGSLTLGLFRINESAGEIIDIINIAKIGLHDLRKTTIIPQDPVLF 1380
Qy 1381 SGSLRMNLDPPSOYSDDEWMTSLAHLKDPVSAIPDKLHECAGEGENUSVGOVOLVCL 1440
Db 1381 SGSLRMNLDPPSOYSDDEWMTSLAHLKDPVSAIPDKLHECAGEGENUSVGOVOLVCL 1440
Qy 1441 ARALLRKTKIIVLDEATAVVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVVL 1500
Db 1441 ARALLRKTKIIVLDEATAVVDLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRVVL 1500
Qy 1501 DKGEIOEYGAPSDILLQORGLFYSMAKDAGLV 1531
Db 1501 DKGEIOEYGAPSDILLQORGLFYSMAKDAGLV 1531

RESULT 10
US-10-756-149-5033
; Sequence 5033, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Zichnik, Albet
; APPLICANT: Zichnik, Albet
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: file
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5033
; LENGTH: 1531
; TYPE: PRF
; ORGANISM: Homo Sapiens
US-10-756-149-5033

Query Match 78.5%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRGFCGADGSDPLMDMNTWNTSNDFPKCFQNTVLYWVPCYLMACFPFFYLYLSRH 60
Db 1 MALRGFCGADGSDPLMDMNTWNTSNDFPKCFQNTVLYWVPCYLMACFPFFYLYLSRH 60
Qy 61 DRGYIOMPLNKTTATGFLIMIVCMADLFVSPMERSGIFLAVPLVSPTLGTTTLA 120
Db 61 DRGYIOMPLNKTTATGFLIMIVCMADLFVSPMERSGIFLAVPLVSPTLGTTTLA 120
Qy 121 TFLQLERRKGVQSSGIMLTFMLVALVCALAILRSKIMTALKEDAOVDLFFRDIFFYVFS 180
Db 121 TFLQLERRKGVQSSGIMLTFMLVALVCALAILRSKIMTALKEDAOVDLFFRDIFFYVFS 180
Qy 181 LLLIQLVLSGSDSPLESETIHDNPNCPSSASFLSITFMWITGLIVRGROPLEGSD 240
Db 181 LLLIQLVLSGSDSPLESETIHDNPNCPSSASFLSITFMWITGLIVRGROPLEGSD 240
Qy 241 LMSLNKEDTSQOVVPLVKNMKKECAKTRKQPVKVYVSSKDPAPCKESSKYDANEVEAL 300
Db 241 LMSLNKEDTSQOVVPLVKNMKKECAKTRKQPVKVYVSSKDPAPCKESSKYDANEVEAL 300
Qy 301 IVKSPQKEMNPSLFKVLYKTGPFYFLMSFPFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360
Db 301 IVKSPQKEMNPSLFKVLYKTGPFYFLMSFPFKAIHDLMMFSGPQILKLLIKFVNDTKAPD 360
Qy 361 WQGFYTYLLFVTACLOTLVLHQYFHI CFVSGMR IKTAIVIGAVYRKALVITNSARKSSTV 420
Db 361 WQGFYTYLLFVTACLOTLVLHQYFHI CFVSGMR IKTAIVIGAVYRKALVITNSARKSSTV 420
```

Db 361 WQGYFTVLLFTVLAQTLVHQYFHCFSWGMRKTAIVAGVYRKALVITNSARKSTV 420
Qy 421 GEIYNLMSVDAORFMDLATYTNMTWSAPLOYTLALYTLNLGSGSVLAGVAVMTLMPVN 480
Db 421 GEIYNLMSVDAORFMDLATYTNMTWSAPLOYTLALYTLNLGSGSVLAGVAVMTLMPVN 480
Qy 481 AVMAKTKTYOVAMKSKDNRIKLMNEILNGIKYIKLYAMELAFKDKVLAIRQBELKYLK 540
Db 481 AVMAKTKTYOVAMKSKDNRIKLMNEILNGIKYIKLYAMELAFKDKVLAIRQBELKYLK 540
Qy 541 KSAVLASVGTFTWCTPPLVALCTFPAVYVTIDENNILDAQTAFAVSLAFNIRPPLNLP 600
Db 541 KSAVLASVGTFTWCTPPLVALCTFPAVYVTIDENNILDAQTAFAVSLAFNIRPPLNLP 600
Qy 601 MVTSSIVQASVSLKRLIFLSHELLEPDSIFRRPVKGGGNGSTVYNNATFTTARSDDPT 660
Db 601 MVTSSIVQASVSLKRLIFLSHELLEPDSIFRRPVKGGGNGSTVYNNATFTTARSDDPT 660
Qy 661 LNCITFESIPEGALVAVVQVCGSGSLLSALLAEMDKVEGHVAKGSVAVYPOQAMQND 720
Db 661 LNCITFESIPEGALVAVVQVCGSGSLLSALLAEMDKVEGHVAKGSVAVYPOQAMQND 720
Qy 721 SLRENILFGCQLEBPYRSVYIQAACALLPDLLEILPSGDRTEIGEKVNLSSGQKQVSLAR 780
Db 721 SLRENILFGCQLEBPYRSVYIQAACALLPDLLEILPSGDRTEIGEKVNLSSGQKQVSLAR 780
Qy 781 AVYSNADIIYLPDPLSLAVDAVGHKIFBNVIGPYGMKLNKTRILVTHSMSTLPQVDYIIV 840
Db 781 AVYSNADIIYLPDPLSLAVDAVGHKIFBNVIGPYGMKLNKTRILVTHSMSTLPQVDYIIV 840
Qy 841 MSGGKISEMGSYOBLARDPAPAEFLRTYASTEOEDABENGVTVSGPGGEAKOMNGM 900
Db 841 MSGGKISEMGSYOBLARDPAPAEFLRTYASTEOEDABENGVTVSGPGGEAKOMNGM 900
Qy 901 LVTDASQKOLQRLSSSSSYSGDISRHHNSTAELOKAEKKEETWKLMEADKQGTQVKL 960
Db 901 LVTDASQKOLQRLSSSSSYSGDISRHHNSTAELOKAEKKEETWKLMEADKQGTQVKL 960
Qy 961 SVYWDYKAIAGLFSPLSIFLPMGNHVSAASNYWLSMTDDPIVNGTOEHTKRLSVYG 1020
Db 961 SVYWDYKAIAGLFSPLSIFLPMGNHVSAASNYWLSMTDDPIVNGTOEHTKRLSVYG 1020
Qy 1021 ALGISOGIAVFGVMAVSIGSILASRCLHYDLHSLIRSPMSFEERPSGNLVNRSKEL 1080
Db 1021 ALGISOGIAVFGVMAVSIGSILASRCLHYDLHSLIRSPMSFEERPSGNLVNRSKEL 1080
Qy 1081 DTVDMSIPEYIKMFMGSLFNVIIGACIYILATPIAIIIPPLGLIYFVQRFYVASSROL 1140
Db 1081 DTVDMSIPEYIKMFMGSLFNVIIGACIYILATPIAIIIPPLGLIYFVQRFYVASSROL 1140
Qy 1141 KRLSVSRSPVYSHFNETLLGVSIVIRAPBEOERFIHQSDLKVDENQKAYPSIVANWLA 1200
Db 1141 KRLSVSRSPVYSHFNETLLGVSIVIRAPBEOERFIHQSDLKVDENQKAYPSIVANWLA 1200
Qy 1201 VRLCEVNCNIVLPALPAVVISRHSLSAGLVGSYSLOYTYTNMIVRMSSEMETNIVA 1260
Db 1201 VRLCEVNCNIVLPALPAVVISRHSLSAGLVGSYSLOYTYTNMIVRMSSEMETNIVA 1260
Qy 1261 VERLKEVSETEKAPMOIQETAPPSWPOVGRVFRNYCLRYREDDLFVLRHINVTNGG 1320
Db 1261 VERLKEVSETEKAPMOIQETAPPSWPOVGRVFRNYCLRYREDDLFVLRHINVTNGG 1320
Qy 1321 EKVGIVGRTAGAGSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
Db 1321 EKVGIVGRTAGAGSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
Qy 1381 SSGLRNMLDBPFOYSDEEWTSLSLAHLKDFVSLPKLDBHECGEGENISVGRQOLVCL 1440
Db 1381 SSGLRNMLDBPFOYSDEEWTSLSLAHLKDFVSLPKLDBHECGEGENISVGRQOLVCL 1440
Qy 1441 ARALLRKTKTILVLEDEATAVADLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRYIVL 1500
Db 1441 ARALLRKTKTILVLEDEATAVADLETDLLIOSTIRTOFEDCTVLTIAHRLNTIMDYTRYIVL 1500

Qy 1501 DKGEIOEYGA PSDLQORGLFYSAKADAGLV 1531
Db 1501 DKGEIOEYGA PSDLQORGLFYSAKADAGLV 1531
RESULT 11
US-09-939-853A-87
; Sequence 87, Application US/09939853A
; Publication No. US20040039163A1
; GENERAL INFORMATION:
; APPLICANT: Burgess et al.
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-099
; CURRENT APPLICATION NUMBER: US/09/939, 853A
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,191
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: 60/267,300
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/269,961
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/277,337
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 1515
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-853A-87
Query Match 77.6%; Score 7769; DB 3; Length 1515;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 17 DMVVTNNTSPDPTKCGQNTVWVWPCFYLMACPPFYFLSLSHDRGYIOMTPLNKTCTA 76
Db 1 DMVVTNNTSPDPTKCGQNTVWVWPCFYLMACPPFYFLSLSHDRGYIOMTPLNKTCTA 60
Qy 77 LGEFLMIVCADLIFYSWERSRGIPLAPVPLVSGPTLLGTTLATFLIQERRKGVSSG 136
Db 61 LGEFLMIVCADLIFYSWERSRGIPLAPVPLVSGPTLLGTTLATFLIQERRKGVSSG 120
Qy 137 IMTFLVALVCAIILRSKIMTALKEDAQVDLFRDITFYVYFSLLIQVLSCFSRSP 196
Db 121 IMTFLVALVCAIILRSKIMTALKEDAQVDLFRDITFYVYFSLLIQVLSCFSRSP 180
Qy 197 LFSETHDRPCRESSASFLSRTTFPMWITGLIVRGYRQPLEGSDLSLNKEDTSQVVPV 256
Db 181 LFSETHDRPCRESSASFLSRTTFPMWITGLIVRGYRQPLEGSDLSLNKEDTSQVVPV 240
Qy 257 LVNWKKECKAKTKOPKVVYSSKDPAPQKSSKVDANEVEALIVKSPKEMNPISLFKV 316
Db 241 LVNWKKECKAKTKOPKVVYSSKDPAPQKSSKVDANEVEALIVKSPKEMNPISLFKV 300
Qy 317 LYKTFGPFLMSFFFAIHDLMFSGPQILKLIKFNVDTKADWQGYFTVLLFPYACL 376
Db 301 LYKTFGPFLMSFFFAIHDLMFSGPQILKLIKFNVDTKADWQGYFTVLLFPYACL 360
Qy 377 QTLVHLQYFHICVSGMRKITAIVAGVYRKALVITNSARSSIVGEIVNLMSVDAORFMD 436
Db 361 QTLVHLQYFHICVSGMRKITAIVAGVYRKALVITNSARSSIVGEIVNLMSVDAORFMD 420
Qy 437 LATYINMISAPQVILALYTLMLNGPSVLAGVAVVLMVFNANVAMAKTKTYOVAAHK 496
Db 421 LATYINMISAPQVILALYTLMLNGPSVLAGVAVVLMVFNANVAMAKTKTYOVAAHK 480
Qy 497 SKDNRIKLMNEILNGIKYIKLYAMELAFKDKVLAIRQBELKYLKKSAYISAVGTFTWCT 556
Db 481 SKDNRIKLMNEILNGIKYIKLYAMELAFKDKVLAIRQBELKYLKKSAYISAVGTFTWCT 540
Qy 557 PFLVALCTPAVYVTIDENNILDAQTAFAVSLAFNIRPPLNLPMTSSIVQASVSLKRL 616

```
|||||
Db      541 PFLVALCTFAVYVTTIDENNILDAQTAFAVSLAFNLRPEPLILPMVSISSIVQASVSKRL 600
Qy      617 RIFLSHELEPDSIERRPVVDGGGNSITVNAFTWARSPPPLNGLTSPISPGALVAV 676
Db      601 RIFLSHELEPDSIERRPVVDGGGNSITVNAFTWARSPPPLNGLTSPISPGALVAV 660
Qy      677 VGQVCGCKSSLSALLAEMDKVEGHVAIKGSVAVYPOQAWIQNDSLRENILFGQLEBPY 736
Db      661 VGQVCGCKSSLSALLAEMDKVEGHVAIKGSVAVYPOQAWIQNDSLRENILFGQLEBPY 720
Qy      737 YRSVIQACALLPDLIELPSGDRTEIGEGKGNVLSGGQKRVSLARAVYSNADIIYFDPLS 796
Db      721 YRSVIQACALLPDLIELPSGDRTEIGEGKGNVLSGGQKRVSLARAVYSNADIIYFDPLS 780
Qy      797 AYDAHVGHKIPENYIGPKGMKNKTRILVTHSMSTLPQVDYIIVMSGKISBMSYOBLL 856
Db      781 AYDAHVGHKIPENYIGPKGMKNKTRILVTHSMSTLPQVDYIIVMSGKISBMSYOBLL 840
Qy      857 ARDGAFAFLRTVASTEOBODAEENGVTGVSQPKBEAKOMENGMVYDTSAGKOLQROLSS 916
Db      841 ARDGAFAFLRTVASTEOBODAEENGVTGVSQPKBEAKOMENGMVYDTSAGKOLQROLSS 900
Qy      917 SSSYSGDISRHNSTAELOKAEAKKEETWKLMADKAOTGQVKSSTYWDYKAIIGLPTSF 976
Db      901 SSSYSGDISRHNSTAELOKAEAKKEETWKLMADKAOTGQVKSSTYWDYKAIIGLPTSF 960
Qy      977 LSIFELFMCHNYSALASNYWLSLMTDDPIVNGIOETHKRLSVYGAALGISOIAVFGYSMA 1036
Db      961 LSIFELFMCHNYSALASNYWLSLMTDDPIVNGIOETHKRLSVYGAALGISOIAVFGYSMA 1020
Qy      1037 VSIGGILASRCLAHVDLHSLRSFMSFPERTPSGNLVNRFSKELDTVDMSIPEVIKMEFG 1096
Db      1021 VSIGGILASRCLAHVDLHSLRSFMSFPERTPSGNLVNRFSKELDTVDMSIPEVIKMEFG 1080
Qy      1097 SLENVIGACIYIILATPIAIIIPPLGIYFFVQGFYASSRQULKLESVSRSPYSHFN 1156
Db      1081 SLENVIGACIYIILATPIAIIIPPLGIYFFVQGFYASSRQULKLESVSRSPYSHFN 1140
Qy      1157 ETLIGVSVIRAFEBEORFIHOSDLKVDENOKAYPSIYANRWLAVRLCVCNCTIVLPAAL 1216
Db      1141 ETLIGVSVIRAFEBEORFIHOSDLKVDENOKAYPSIYANRWLAVRLCVCNCTIVLPAAL 1200
Qy      1217 FAVISRHSLSAGLVGLSVYSILOVTTYLNMVLVRMSSEMETNIIVAVERLKEYSETEKAPW 1276
Db      1201 FAVISRHSLSAGLVGLSVYSILOVTTYLNMVLVRMSSEMETNIIVAVERLKEYSETEKAPW 1260
Qy      1277 QIOETAPSSWPQVGRVFEFRNYCLARYEDLDVFLRHINVTINGEKVGI VGRTOGAKSSL 1336
Db      1261 QIOETAPSSWPQVGRVFEFRNYCLARYEDLDVFLRHINVTINGEKVGI VGRTOGAKSSL 1320
Qy      1337 TLGFRIINESAGEIIIDIGINIAKIGLHDLRFKITIIPQDPVLPSSGSLRMNLDPPSOQSD 1396
Db      1321 TLGFRIINESAGEIIIDIGINIAKIGLHDLRFKITIIPQDPVLPSSGSLRMNLDPPSOQSD 1380
Qy      1397 BEVWTSLELAHKDFVSALPKLDHECAGGSENLISVQOROLVCLARALLRKTILVILDEA 1456
Db      1381 BEVWTSLELAHKDFVSALPKLDHECAGGSENLISVQOROLVCLARALLRKTILVILDEA 1440
Qy      1457 TAAVDELTDLLOSTIRTOFEDCTVLTIAHRLNTIMDYTRYIVILDKGEIOEGAPSDILQ 1516
Db      1441 TAAVDELTDLLOSTIRTOFEDCTVLTIAHRLNTIMDYTRYIVILDKGEIOEGAPSDILQ 1500
Qy      1517 ORGLFYSAKADAGLV 1531
Db      1501 ORGLFYSAKADAGLV 1515
```

```
APPLICANT: Day, Anthony G.
APPLICANT: Estell, David A.
APPLICANT: Lyons, Eric H.
APPLICANT: Yao, Jian
TITLE OR INVENTION: Methods for Modulating Proteins Not
FILE REFERENCE: GC773-2
CURRENT APPLICATION NUMBER: US/10/618, 281
PRIOR FILING DATE: 2003-07-11
PRIOR APPLICATION NUMBER: US 60/395,325
NUMBER OF SEQ ID NOS: 92
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 42
LENGTH: 1515
TYPE: PR
ORGANISM: Homo sapiens
US-10-618-281-42
```

Query Match 77.5%; Score 7763; DB 5; Length 1515;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1514; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      17 DNANTNTSNPDFTKCHQNTVLVWPCFYLMACPFYFLYSRDRGTYQMTPLNKTITA 76
Db      1 DNANTNTSNPDFTKCHQNTVLVWPCFYLMACPFYFLYSRDRGYIQMTPLNKTITA 60
Qy      77 LGFLLMIVCANADLYFSWERSRGIFLAPVFLVSPTLIGITTLATPFIQLERRGVQSSG 136
Db      61 LGFLLMIVCANADLYFSWERSRGIFLAPVFLVSPTLIGITTLATPFIQLERRGVQSSG 120
Qy      137 IMLEFWVALVICALAIRSKIMTALKEDAQVDLFRDITFYVYFSLILQVLSCFSRSP 196
Db      121 IMLEFWVALVICALAIRSKIMTALKEDAQVDLFRDITFYVYFSLILQVLSCFSRSP 180
Qy      197 LFSFTIHDPNCPBSSASFSRITFWMITGLIYRGYQPLBGSJLMSLNKEDTSEQVVPV 256
Db      181 LFSFTIHDPNCPBSSASFSRITFWMITGLIYRGYQPLBGSJLMSLNKEDTSEQVVPV 240
Qy      257 LVKWKKECAKTRQAPKVVVYSSKOPAPKSSKVDNNEFEALIVSPKEMNPSLKV 316
Db      241 LVKWKKECAKTRQAPKVVVYSSKOPAPKSSKVDNNEFEALIVSPKEMNPSLKV 300
Qy      317 LYKTFGYPFLMSFFFKAIHDLMMFSGPQILKLIKFNVDTKAPDMQGYFTVLLFTVACL 376
Db      301 LYKTFGYPFLMSFFFKAIHDLMMFSGPQILKLIKFNVDTKAPDMQGYFTVLLFTVACL 360
Qy      377 QTLVLAHQYFHI CFVSGNRITKTAIVGAYYKALVITNSARKSSTVGEIVNLMVSUQRFMD 436
Db      361 QTLVLAHQYFHI CFVSGNRITKTAIVGAYYKALVITNSARKSSTVGEIVNLMVSUQRFMD 420
Qy      437 LATYNNIWSAPLOVITLALVLMNLGPSYLAGAVWVWLVNVAWNAWMTKTKYQVAHMK 496
Db      421 LATYNNIWSAPLOVITLALVLMNLGPSYLAGAVWVWLVNVAWNAWMTKTKYQVAHMK 480
Qy      497 SKDNRIKLMEIILNGIVLKLAYAMELAFKQVLAIROBELKVLKKSAYLSAVGTFTWCT 556
Db      481 SKDNRIKLMEIILNGIVLKLAYAMELAFKQVLAIROBELKVLKKSAYLSAVGTFTWCT 540
Qy      557 PFLVALCTFAVYVTTIDENNILDAQTAFAVSLAFNLRPEPLILPMVSISSIVQASVSKRL 616
Db      541 PFLVALCTFAVYVTTIDENNILDAQTAFAVSLAFNLRPEPLILPMVSISSIVQASVSKRL 600
Qy      617 RIFLSHELEPDSIERRPVVDGGGNSITVNAFTWARSPPPLNGLTSPISPGALVAV 676
Db      601 RIFLSHELEPDSIERRPVVDGGGNSITVNAFTWARSPPPLNGLTSPISPGALVAV 660
Qy      677 VGQVCGCKSSLSALLAEMDKVEGHVAIKGSVAVYPOQAWIQNDSLRENILFGQLEBPY 736
Db      661 VGQVCGCKSSLSALLAEMDKVEGHVAIKGSVAVYPOQAWIQNDSLRENILFGQLEBPY 720
Qy      737 YRSVIQACALLPDLIELPSGDRTEIGEGKGNVLSGGQKRVSLARAVYSNADIIYFDPLS 796
```

```

Db      721 YRSVIAQACALLPDEILIPSGDRTEIGKGNVLGGQKQVSLAARAVSNADIVLPDDPLS 780
Qy      797 AVDAHVGKHIPEENVYIGPKGMLKNKTRILIVTHSMSTLPQVDTIIVMSGKISMGSYOELL 856
Db      781 AVDAHVGKHIPEENVYIGPKGMLKNKTRILIVTHSMSTLPQVDTIIVMSGKISMGSYOELL 840
Qy      857 ARDGAFAEFLRTYASTEOBDAEENGVTGVSQPGKEAKQOMENGLVTDTSAGKOJLOROLSS 916
Db      841 ARDGAFAEFLRTYASTEOBDAEENGVTGVSQPGKEAKQOMENGLVTDTSAGKOJLOROLSS 900
Qy      917 SSSYSGDISRHHNSTAELOKAEAKKEETWKLMEADKAQOTGVKLSVVDYMKALIGLFISSF 976
Db      901 SSSYSGDISRHHNSTAELOKAEAKKEETWKLMEADKAQOTGVKLSVVDYMKALIGLFISSF 960
Qy      977 LSTLFMCNHNVSALASNTWLSMTDDPIVNGTOEHTKRLSVYGALGISQGIANFGYSMA 1036
Db      961 LSTLFMCNHNVSALASNTWLSMTDDPIVNGTOEHTKRLSVYGALGISQGIANFGYSMA 1020
Qy      1037 VSIGGILASRCLHVDLHSLIRSPMSFPERTPSGNLVNRFSGKLDVDSMTPEVTKMFMG 1096
Db      1021 VSIGGILASRCLHVDLHSLIRSPMSFPERTPSGNLVNRFSGKLDVDSMTPEVTKMFMG 1080
Qy      1097 SLFNVIGACIVILLATPIAIIIPPLGLIYFVQRFYVASSRQKRLSEVSRSPPYSHFN 1156
Db      1081 SLFNVIGACIVILLATPIAIIIPPLGLIYFVQRFYVASSRQKRLSEVSRSPPYSHFN 1140
Qy      1157 ETLIGSVYIRAFEEOEERFIHOSDLKVDENQKAYPSIVANRWLAVRLECVGNCTIVLFAAL 1216
Db      1141 ETLIGSVYIRAFEEOEERFIHOSDLKVDENQKAYPSIVANRWLAVRLECVGNCTIVLFAAL 1200
Qy      1217 FAVISRSLSAGLVGLSVSYSLQVTTYTNMLVRMSSEMETNIIVVERLKEYSSETEKAPW 1276
Db      1201 FAVISRSLSAGLVGLSVSYSLQVTTYTNMLVRMSSEMETNIIVVERLKEYSSETEKAPW 1260
Qy      1277 QIOETAPSSWPQYGRVEFRNYCYRREDLDFVLRIHIVTINGEKVYIGVRTAGKSSL 1336
Db      1261 QIOETAPSSWPQYGRVEFRNYCYRREDLDFVLRIHIVTINGEKVYIGVRTAGKSSL 1320
Qy      1337 TLGLFRINESAGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLFSGSLRMNLDPFSQYSD 1396
Db      1321 TLGLFRINESAGEIIIDGINIAKIGLHDLRFKTTIIPQDPVLFSGSLRMNLDPFSQYSD 1380
Qy      1397 BEWVTSLELAHLKQFVSALPKLDEHCEGEGENISVQOROUVCARALARTKILVDEA 1456
Db      1381 BEWVTSLELAHLKQFVSALPKLDEHCEGEGENISVQOROUVCARALARTKILVDEA 1440
Qy      1457 TAAVDLETDDLIOSTIRTOFEDCTVLTAHRLNTIMDYTRIVYLDKGBIOEYGAPSDLLQ 1516
Db      1441 TAAVDLETDDLIOSTIRTOFEDCTVLTAHRLNTIMDYTRIVYLDKGBIOEYGAPSDLLQ 1500
Qy      1517 QRGLFYSMAKDAGLV 1531
Db      1501 QRGLFYSMAKDAGLV 1515

```

```

RESULT 13
US-10-408-765A-1718
; Sequence 1718, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fany, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Marnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408, 765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1718
; LENGTH: 1388
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-408-765A-1718

Query Match      70.3%; Score 7037; DB 4; Length 1388;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 1387; Conservative 0; Mismatches 1; Indels 42; Gaps 1;

Qy      67 MTPLNKTKTALGFLMTIVCMADLFYSFMEBSRGIFLAPVPLVSPTLGITTLLATPLIQ 126
Db      1 MTPLNKTKTALGFLMTIVCMADLFYSFMEBSRGIFLAPVPLVSPTLGITTLLATPLIQ 60
Qy      127 ERRKVGSSGIMLTFMWLVALCALAIRSKIMTAKEDAQVLDPRDITFYVYFSLILQL 186
Db      61 ERRKVGSSGIMLTFMWLVALCALAIRSKIMTAKEDAQVLDPRDITFYVYFSLILQL 97
Qy      187 VLSCFSRSLPFSETHIDPNPCPSSASFLSRITFMWITGLIYNGYQPLEGSLMSLNK 246
Db      98 -----NCPSSASFLSRITFMWITGLIYNGYQPLEGSLMSLNK 138
Qy      247 EDTSEQVVPVLVKNWKKCAKTRQPKYVYSSKDPQPKSSKVDANBEVEALIVKSPQ 306
Db      139 EDTSEQVVPVLVKNWKKCAKTRQPKYVYSSKDPQPKSSKVDANBEVEALIVKSPQ 198
Qy      307 KENNPSLEKVLKTKGFPYFLMSFFPKAIDHLMFSGPQILKLIKFNNDTKAPDMQGYFY 366
Db      199 KENNPSLEKVLKTKGFPYFLMSFFPKAIDHLMFSGPQILKLIKFNNDTKAPDMQGYFY 258
Qy      367 TVLLFTVACIQTLVHGYFHI CFVSGMRITAVTGAAYRKALVTTNSARKSSTYGEIYNL 426
Db      259 TVLLFTVACIQTLVHGYFHI CFVSGMRITAVTGAAYRKALVTTNSARKSSTYGEIYNL 318
Qy      427 MSVDAQRFMDLATYINNIWASAPLOVITLALYLMNTLPSVLAGAVANVWVNAVWAMK 486
Db      319 MSVDAQRFMDLATYINNIWASAPLOVITLALYLMNTLPSVLAGAVANVWVNAVWAMK 378
Qy      487 TKTYQVAMHMSKONRIKLANEIIINGIKVLKLYAMELAFKQKVLAIROBELKVLKKSAYLS 546
Db      379 TKTYQVAMHMSKONRIKLANEIIINGIKVLKLYAMELAFKQKVLAIROBELKVLKKSAYLS 438
Qy      547 AVGTFTWCTPPELVALCTFAVYVTTIDENNILDQAFAVSLAFNIIIRPPLNIIIPMVISI 606
Db      439 AVGTFTWCTPPELVALCTFAVYVTTIDENNILDQAFAVSLAFNIIIRPPLNIIIPMVISI 498
Qy      607 VQASVSLKRLRIFLSHELEPPDSIERRPVVDGGGTSITVRNATFTWARSDDPTLNGITF 666
Db      499 VQASVSLKRLRIFLSHELEPPDSIERRPVVDGGGTSITVRNATFTWARSDDPTLNGITF 558
Qy      667 SIPEGALVAVGVGCGKSSLSALLAEMDKVEGHVAKISVAVVPOQAMIONDSLRENI 726
Db      559 SIPEGALVAVGVGCGKSSLSALLAEMDKVEGHVAKISVAVVPOQAMIONDSLRENI 618
Qy      727 LFGQULEEPPYRSYVQACALLPDEILIPSGDRTEIGKGNVLGGQKQVSLAARAVSNA 786
Db      619 LFGQULEEPPYRSYVQACALLPDEILIPSGDRTEIGKGNVLGGQKQVSLAARAVSNA 678
Qy      787 DYLFPDPLSAVDHVGKHI FENVIGPKGMLKNKTRILIVTHSMSTLPQVDTIIVMSGKI 846
Db      679 DYLFPDPLSAVDHVGKHI FENVIGPKGMLKNKTRILIVTHSMSTLPQVDTIIVMSGKI 738
Qy      847 SEMGSYQELLARDAFAEFLRTYASTEOBDAEENGVTGVSQPGKEAKQOMENGLVYDSA 906
Db      739 SEMGSYQELLARDAFAEFLRTYASTEOBDAEENGVTGVSQPGKEAKQOMENGLVYDSA 798
Qy      907 GKOLOROLSSSSSGDISRHHNSTAELOKAEAKKEETWKLMEADKAQOTGVKLSVVDY 966
Db      799 GKOLOROLSSSSSGDISRHHNSTAELOKAEAKKEETWKLMEADKAQOTGVKLSVVDY 858
Qy      967 MKAIGLFISSLIFLFCNHNVSALASNTWLSMTDDPIVNGTOEHTKRLSVYGALGISQ 1026

```

Db 859 MKAIGLFISFLSIFLPMCNHVSALASNYWLSLMTDDPIVNGTQEHTKVRLSVYGAIGISQ 918
Qy 1027 GIAVFGYSMAVSIIGGILASRCILHVDLHLSILRSPMSFFERTPSGNLVNRFSEKELDTVDSM 1086
Db 919 GIAVFGYSMAVSIIGGILASRCILHVDLHLSILRSPMSFFERTPSGNLVNRFSEKELDTVDSM 978
Qy 1087 IPEVITKPMGSLFNVIIGACIYIILATPIAAIIPPLGLIYFFVQRFYVASSRQLKRLSEV 1146
Db 979 IPEVITKPMGSLFNVIIGACIYIILATPIAAIIPPLGLIYFFVQRFYVASSRQLKRLSEV 1038
Qy 1147 SRSVYSHFNFTLLGVSVIRAFEEORERFIHOSDLKVDENOKAYYSIVANMILAVRLECV 1206
Db 1039 SRSVYSHFNFTLLGVSVIRAFEEORERFIHOSDLKVDENOKAYYSIVANMILAVRLECV 1098
Qy 1207 GNCIVLFAALFAVISRHSLSAGLVGLSVSYSLOVTTYLNMLVRMSSEMETNIIVAVERLKE 1266
Db 1099 GNCIVLFAALFAVISRHSLSAGLVGLSVSYSLOVTTYLNMLVRMSSEMETNIIVAVERLKE 1158
Qy 1267 YSETEKEAPMOIOETAPPSNPQVGRVFRNVCILRYREDDLPVLRIHNVTTINGEKVGI 1326
Db 1159 YSETEKEAPMOIOETAPPSNPQVGRVFRNVCILRYREDDLPVLRIHNVTTINGEKVGI 1218
Qy 1327 GRTAGKSSLTGLFRINESABGRIIDGINIAKTGLHDLRPKTIIIPDPLVPSGSLRM 1386
Db 1219 GRTAGKSSLTGLFRINESABGRIIDGINIAKTGLHDLRPKTIIIPDPLVPSGSLRM 1278
Qy 1387 NLDPFSQYSDDEWVTSLELAHLKDFVSALPKLDHECAEGENLSVGOROLVCLARALLR 1446
Db 1279 NLDPFSQYSDDEWVTSLELAHLKDFVSALPKLDHECAEGENLSVGOROLVCLARALLR 1338
Qy 1447 KTKILVLDEATPAVDLETDLIQSTIRIQFEDCTVLTIAHRLNTIMDYTR 1496
Db 1339 KTKILVLDEATPAVDLETDLIQSTIRIQFEDCTVLTIAHRLNTIMDYTR 1388

RESULT 14

US-09-939-853A-88
Sequence 88, Application US/09939853A
Publication No. US20040039163A1
GENERAL INFORMATION:
APPLICANT: Burgess et al.
TITLE OF INVENTION: No. US20040039163A1 Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-099
CURRENT APPLICATION NUMBER: US/09/939, 853A
CURRENT FILING DATE: 2001-08-27
PRIOR APPLICATION NUMBER: 60/228,191
PRIOR FILING DATE: 2000-08-25
PRIOR APPLICATION NUMBER: 60/267,300
PRIOR FILING DATE: 2001-02-08
PRIOR APPLICATION NUMBER: 60/269,961
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/277,337
NUMBER OF SEQ ID NOS: 159
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 88
LENGTH: 1528
TYPE: PRT
ORGANISM: Mus musculus
US-09-939-853A-88

Query Match 69.9%; Score 7002.5; DB 3; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

Qy 1 MALNGFCGADSGDPLMDNNTWNTNSDPFKCFONTVLVWPCEFYLAACPFYFYLSRH 60
Db 1 MALNSFCGADSGDPLMDNNTWNTNSDPFKCFONTVLVWPCEFYLAACPFYFYLSRH 60
Qy 61 DRGIOMPLUNKTKTALGFLLMIWCMADLFYSFWRSHRGIFLAPVFLVSPITLIGITTTLLA 120
Db 61 DRGIOMTHLUNKTKTALGFLLMIWCMADLFYSFWRSHRGIFLAPVFLVSPITLIGITTTLLA 120

Qy 121 TFLIQLERRKGVSSGIMLTFWLVALVCAIALIRSKITMTAKEDAOVDLEFDITFPYFYS 180
Db 121 TFLIQLERRKGVSSGIMLTFWLVALVCAIALIRSKITMTAKEDAOVDLEFDITFPYFYS 180
Qy 181 LLLIQVLSCGSDSPILFSETIHDNPPCPRESSAFSLRITFWMTGLIVRGYRQPLESSD 240
Db 181 LVLQVLVSCGSDSPILFSETIHDNPPCPRESSAFSLRITFWMTGLIVRGYRQPLESSD 240
Qy 241 LMSLNKEDTSEQVVPVLVWKKKECAKTRKQPVNVYVS-SKDPAPQKSSKVDAENEVEYA 299
Db 241 LMSLNKEDTSEQVVPVLVWKKKECAKTRKQPVNVYVS-SKDPAPQKSSKVDAENEVEYA 300
Qy 300 LIVSPQKMNPSLPKVLKTFPGPYFLMSFFPKAIDHLMFSGPQILKLIKPVNDIYAP 359
Db 301 LIVSPHKBDEPSPFLKVLKTFPGPYFLMSFFPKAIDHLMFSGPQILKLIKPVNDIYAP 360
Qy 360 DMQGYFTVLLFVYACIQTALHGYFPHCFVSGMRITAVYGAVYRALKVLTNARSST 419
Db 361 DMQGYFTVLLFVYACIQTALHGYFPHCFVSGMRITAVYGAVYRALKVLTNARSST 420
Qy 420 VGEIYNLMSVDAQRFMDLATYNNMWSAPLOVILATYLNMLNLPVSLAGVAVVLMVVP 479
Db 421 VGEIYNLMSVDAQRFMDLATYNNMWSAPLOVILATYLNMLNLPVSLAGVAVVLMVVP 480
Qy 480 NAYNAMKTKTYQVAHMSKONRIKLAMBIINGIVLKLAMELAPKOKVLAIRQBELKVL 539
Db 481 NAYNAMKTKTYQVAHMSKONRIKLAMBIINGIVLKLAMELAPKOKVMSIROBELKVL 540
Qy 540 KKSAYLAVGFTFWCPELVALCTPAVYVTTIDENNITLDQTAVSLAFPIITLFPNITL 599
Db 541 KKSAYLAVGFTFWCPELVALCTPAVYVTTIDENNITLDQTAVSLAFPIITLFPNITL 600
Qy 600 PMVSIIVQASVSLKRLRILFSLHELEPDSIERSPVKGCGTNSITYRANAFWTARSDDP 659
Db 601 PMVSIIVQASVSLKRLRILFSLHELEPDSIERSPVKGCGTNSITYRANAFWTARSDDP 659
Qy 660 TLNGITFSIPBGALVAVGVGCGKSSLSALAEKMDKVBGHVAKGSVAVYPOQAWTON 719
Db 660 TLNGITFSIPBGALVAVGVGCGKSSLSALAEKMDKVBGHVAKGSVAVYPOQAWTON 719
Qy 720 DSLAEKNTLFGQEEPPYRSYIACALLPDELILPBGDRTEIGKGVNLGGCGKORVSLA 779
Db 720 DSLAEKNTLFGQEEPPYRSYIACALLPDELILPBGDRTEIGKGVNLGGCGKORVSLA 779
Qy 780 RAYVSNADIVLFDPLSAVDAAVGHKIFENVIIGKGMKNKTRILVTHSMSSILPOVDYI 839
Db 780 RAYVSNADIVLFDPLSAVDAAVGHKIFENVIIGKGMKNKTRILVTHSMSSILPOVDYI 839
Qy 840 VMSGGKISEMGSYQELLARDGAPAEPLRTYASTQEDQDAEENGVTGVSQPKKAKQWENG 899
Db 840 VMSGGKISEMGSYQELLARDGAPAEPLRTYASTQEDQDAEENGVTGVSQPKKAKQWENG 896
Qy 900 MLVYDSACKQLOROLSSSSSYSGDISRHNSTALQRAAKKETWLMEDAKQOTGVYK 959
Db 900 MLVYDSACKQLOROLSSSSSYSGDISRHNSTALQRAAKKETWLMEDAKQOTGVYK 955
Qy 960 LSVYMDVWKAIGLISFLSIFLPMCNHVSALASNYWLSLMTDD-PIYNGTQEHTKVRLSV 1018
Db 960 LSVYMDVWKAIGLISFLSIFLPMCNHVSALASNYWLSLMTDD-PIYNGTQEHTKVRLSV 1015
Qy 1019 YGALGISQGIADVFGYSMAVSIIGGILASRCILHVDLHLSILRSPMSFFERTPSGNLVNRFSEK 1078
Db 1019 YGALGISQGIADVFGYSMAVSIIGGILASRCILHVDLHLSILRSPMSFFERTPSGNLVNRFSEK 1075
Qy 1079 ELDTVDNMIPEVITKPMGSLFNVIIGACIYIILATPIAAIIPPLGLIYFFVQRFYVASSR 1138
Db 1079 ELDTVDNMIPEVITKPMGSLFNVIIGACIYIILATPIAAIIPPLGLIYFFVQRFYVASSR 1135
Qy 1139 QLRKLESYSRSPVYSHFNFTLLGVSVIRAFEEORERFIHOSDLKVDENOKAYYSIVANRW 1198
Db 1139 QLRKLESYSRSPVYSHFNFTLLGVSVIRAFEEORERFIHOSDLKVDENOKAYYSIVANRW 1195
Qy 1199 LAVRLCEVGNCIYVFAALFAVISRHSLSAGLVGLSVSYSLOVTTYLNMLVRMSSEMETNI 1258

Db 1196 LAARLECVGNCIVLFAALPNTVISHSLSAGLVGSVSLQITAYLMLVLRMSSEMTNI 1255
Qy 1259 VAVBRLEKSEYSETEAEAWMOIQTETPSSWPOVGEVERFRNYCLRRREDLPVLRHINTIN 1318
Db 1256 VAVBRLEKSEYSETEAEAWMOIQTETPSSWPOVGEVERFRNYCLRRREDLPVLRHINTIN 1315
Qy 1319 GGEKVGIVGRITGAGKSSLTGLFPAINSABEIIIDGINIAKIGLHDLRFKTTIIPDDV 1378
Db 1316 GGEKVGIVGRITGAGKSSLTGLFPAINSABEIIIDGINIAKIGLHDLRFKTTIIPDDV 1375
Qy 1379 LFGSGLMNLDPFSQYSDDEWVTSLELAHLNDPVSALPDKLDEHCAEGENTSVGQRLV 1438
Db 1376 LFGSGLMNLDPFSQYSDDEWVTSLELAHLNDPVSALPDKLDEHCAEGENTSVGQRLV 1435
Qy 1439 CLARALLRKTKIIVLDATAVVDLETDDLIQSTIRTOPEDECTVTIARLNTIMDYRVI 1498
Db 1436 CLARALLRKTKIIVLDATAVVDLETDDLIQSTIRTOPEDECTVTIARLNTIMDYRVI 1495
Qy 1499 VLDKGEIOEYGAPEBDLQORGLFYSMAKDGLV 1531
Db 1496 VLDKGEIOEYGAPEBDLQORGLFYSMAKDGLV 1528

RESULT 15
US-10-618-281-63
Sequence 63, Application US/10618281
Publication No. US20040219609A1
GENERAL INFORMATION:
APPLICANT: Day, Anthony G.
APPLICANT: Estell, David A.
APPLICANT: Lyons, Eric H.
APPLICANT: Yao, Jian
TITLE OF INVENTION: Methods for Modulating Proteins Not
FILE REFERENCE: GC773-2
CURRENT APPLICATION NUMBER: US/10/618, 281
PRIORITY FILING DATE: 2003-07-11
PRIORITY APPLICATION NUMBER: US 60/395,325
PRIORITY FILING DATE: 2002-07-12
NUMBER OF SEQ ID NOS: 92
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 63
LENGTH: 1303
TYPE: PRT
ORGANISM: Homo sapiens
US-10-618-281-63

Query Match 60.6%; Score 6068.5; DB 5; Length 1303;
Best Local Similarity 81.4%; Pred. No. 0;
Matches 1229; Conservative 17; Mismatches 57; Indels 207; Gaps 10;

Qy 22 WNTSNPDFTKCFONTVVLVWPCFYLMACPFYFLYLSRHDGYIOMTPLNKTKTALGFL 81
Db 1 WNTSNPDFTKCFONTVVLVWPCFYLMACPFYFLYLSRHDGYIOMTPLNKTKTALGFL 60
Qy 82 WIYCMADLFYSFERSRGIFLAPVFLVPTLLGTTLLATFLIQLERRKGVSSGIMLTF 141
Db 61 WIYCMADLFYSFERSRGIFLAPVFLVPTLLGTTLLATFLIQLERRKGVSSGIMLTF 120
Qy 142 WLVALCALALILSKITAKEDAVNDLFRDIFPVYFSLITQVLVSCSDSPLESET 201
Db 121 WLVALCALALILSKITAKEDAVNDLFRDIFPVYFSLITQVLVSCSDSPLESET 177
Qy 202 IHDPNCPRESSASFLSRITFMWITGLIVRGYROPLESDLSINKEDTSEQVVPVLYKN 261
Db 178 IHDPNCPRESSASFLSRITFMWITGLIVRGYROPLESDLSINKEDTSEQVVPVLYKN 237
Qy 262 KKGCAKTRKQPVKVVYSSKDPADPKSSKYDANEBEVALIVKSPQKEMNPSLFKLYKTF 321
Db 238 KKGCAKTRN-----SSGSGSSCSANTALF-----PA--PTCHKSF 271
Qy 322 GPYFLMSFFFKAIHDLMMFGSPQILKLLIFVNDTKAPDWQGYFYTVLFPVTAQLQTLVL 381

Db 272 QALSL-----LCRLILFVNDTKAPDWQGYFYTVLFPVTAQLQTLVL 314
Qy 382 HOYFHICFVGMRIKTAVICAVYRKALVITNSARKSTVGEIYNLMSVDAQREMDLATYI 441
Db 315 HOYFHICFVGMRIKTAVICAVYRKALVITNSARKSTVGEIYNLMSVDAQREMDLATYI 374
Qy 442 NMTMSAPLOYITLALYLMNLGSPVLAGVAVNMLMPVNVAMMKTTYVAHMKSDNR 501
Db 375 NMTMSAPLOYITLALYLMNLGSPVLAGVAVNMLMPVNVAMMKTTYVAHMKSDNR 416
Qy 502 IKLMNEILNGIKVLKYAMELAFKDKVLAIOBELKYLKKSAYLSAVGTFTWCTPPLVA 561
Db 417 IKLMNEILNGIKVLKYAMELAFKDKVLAIOBELKYLKKSAYLSAVGTFTWCTPPLVA 476
Qy 562 LCTPAVYVITDENNIIDAOCTAFVSLALFNLRPLNLPVVISITVOASVSLKRLIFLS 621
Db 477 LCTPAVYVITDENNIIDAOCTAFVSLALFNLRPLNLPVVISITVOAGE-----A 528
Qy 622 HEELEPDSIERRPVKGCGGINSITVNAFTMARSDPPLNGITFSTPEGALVAVGVG 681
Db 529 GATSEKGPWGSRRKDG-----TRQASFSVABPGVLCRFSITFSLPEGALVAVGVG 581
Qy 682 CGKSSLLSALLAEKMDKVEGHVIAIKGSVAYVPOQAMIONDSLRENILFGCOLLEPYRSVI 741
Db 582 CGKSSLLSALLAEKMDKVEGHVIAIKGSVAYVPOQAMIONDSLRENILFGCOLLEPYRSVI 641
Qy 742 QACALLPDLRLPSGRTETIGKGVNLSSGQKQVSLARAVSNADYLFDDPLSAVDAA 801
Db 642 QACALLPDLRLPSGRTETIGKGVNLSSGQKQVSLARAVSNADYLFDDPLSAVDAA 701
Qy 802 VGHIFBNVIGPQKMLKNTKRIIVYTHSMSTLPVDVYIYVSGKISMSGYQELIARDA 861
Db 702 VGHIFBNVIGPQKMLKNTKRIIVYTHSMSTLPVDVYIYVSGKISMSGYQELIARDA 724
Qy 862 FAEFLRTYASTEQEQAEBEGVTVGSGPKEAKQEMENGLVTDQAKQLQRLSSSSYS 921
Db 725 -----S 725
Qy 922 GDLSRHNSYAELOKAEKKEBTWKLMEADKQOTGVKLSVYDYMKAIGLFISLIFL 981
Db 726 CDL-----QVKLSVYDYMKAIGLFISLIFL 753
Qy 982 FMCNHYSAASNTWLSMTDDPIVNGTOETKRLSYGALGISOGIAVNGYMANYSIG 1041
Db 754 FMCNHYSAASNTWLSMTDDPIVNGTOETKRLSYGALGISOGIAVNGYMANYSIG 813
Qy 1042 ILASRCIADVLHSHIIRSPMSFFERTPSGNLVNRFSEKEDTVDSMTPEVTKMFGSLFNV 1101
Db 814 ILASRCIADVLHSHIIRSPMSFFERTPSGNLVNRFSEKEDTVDSMTPEVTKMFGSLFNV 873
Qy 1102 IGACIVILLATPIAIIIPPLGLIYFVQRFYVASSRQLKRLBSVSRSPYSHNETLIG 1161
Db 874 IGACIVILLATPIAIIIPPLGLIYFVQRFYVASSRQLKRLBSVSRSPYSHNETLIG 933
Qy 1162 VSVIRAFEBQERIHOSDLKVDENQKAYTPIYANWLAARLECVGNCIVLFAALPAVIS 1221
Db 934 VSVIRAFEBQERIHOSDLKVDENQKAYTPIYANWLAARLECVGNCIVLFAALPAVIS 993
Qy 1222 RSHLSAGLVGLSVSYSLQVTTYLNMVLVRMSSEMTNIIVAVBRLEKSEYSETEAEAWMOIQT 1281
Db 994 RSHLSAGLVGLSVSYSLQVTTYLNMVLVRMSSEMTNIIVAVBRLEKSEYSETEAEAWMOIQT 1053
Qy 1282 APPSSWPOVGRVEFRNYCLRRREDLPVLRHINTINGEKVIGVRTGAGKSSLTGLF 1341
Db 1054 APPSSWPOVGRVEFRNYCLRRREDLPVLRHINTINGEKVIGVRTGAGKSSLTGLF 1113
Qy 1342 RINESABEIIIDGINIAKIGLHDLRFKTTIIPDDVLPSSGLRMNLDPFSQYSDDEWVT 1401
Db 1114 RINESABEIIIDGINIAKIGLHDLRFKTTIIPDDVLPSSGLRMNLDPFSQYSDDEWVT 1173
Qy 1402 SLELAHKFVSALPDKLDEHCAEGENTSVGQRLVCLARALLRKTKIIVLDATAVVD 1461

Db	1174	SLELAHLKDFVSALPDKLDHECAEGENLSVGOROLVCLARALLRKTILVLDENTAAND	1233
Qy	1462	LETDDLQSTIRTFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQORGLF	1521
Db	1234	LETDDLQSTIRTFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAPSDLLQORGLF	1293
Qy	1522	YSMAKDAGLV	1531
Db	1294	YSMAKDAGLV	1303

Search completed: December 15, 2005, 15:47:17
Job time : 170.491 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 15:17:40 / Search time 7.34048 Seconds
(without alignments)
1786.344 Million cell updates/sec

Title: US-10-665-283-8

Perfect score: 10016
Sequence: 1 MALRGFCADGSDPLMDMNV.....RSVAVAKKPKFSPDLSL 1947

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 51463 seqs, 6734788 residues

Total number of hits satisfying chosen parameters: 51463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database:
- 1: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/PCR_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2490	24.9	1023	6	US-10-995-561-968 Sequence 968, App
2	2115	21.1	1581	7	US-11-080-439-24 Sequence 24, App1
3	2101.5	21.0	1588	6	US-10-995-561-527 Sequence 527, App
4	2100	21.0	1581	7	US-11-090-439-26 Sequence 26, App1
5	1341.5	13.4	1404	6	US-10-995-561-526 Sequence 526, App
6	969	9.7	410	6	US-10-995-561-969 Sequence 969, App
7	881.5	8.8	407	6	US-10-613-744-14 Sequence 14, App1
8	799	8.0	268	6	US-10-995-561-967 Sequence 967, App
9	788.5	7.9	391	6	US-10-613-744-13 Sequence 13, App1
10	514.5	5.1	548	6	US-10-793-626-326 Sequence 326, App
11	495.5	4.9	622	6	US-10-467-657-1346 Sequence 1346, App
12	468.5	4.7	549	7	US-11-082-389-148 Sequence 148, App
13	419.5	4.2	701	6	US-10-995-561-575 Sequence 575, App
14	419.5	4.2	808	6	US-10-995-561-574 Sequence 574, App
15	419.5	4.2	808	7	US-11-105-268-53 Sequence 53, App1
16	417	4.2	563	7	US-11-055-822-254 Sequence 254, App
17	413.5	4.1	618	6	US-10-467-657-2758 Sequence 2758, App
18	413.5	4.1	618	6	US-10-467-657-6728 Sequence 6728, App
19	322	3.2	538	6	US-10-793-626-260 Sequence 260, App
20	322	3.2	751	6	US-10-995-561-578 Sequence 578, App
21	307.5	3.1	329	6	US-10-995-561-576 Sequence 576, App
22	301	3.0	530	7	US-11-055-822-256 Sequence 256, App1
23	296.5	2.9	564	7	US-11-082-389-78 Sequence 78, App1
24	265	2.6	439	7	US-11-082-389-80 Sequence 80, App1
25	256	2.6	151	7	US-11-082-389-152 Sequence 152, App

26	247.5	2.5	304	7	US-11-082-389-336 Sequence 336, App
27	243.5	2.4	659	6	US-10-995-561-573 Sequence 573, App
28	228	2.3	219	7	US-11-082-389-250 Sequence 250, App
29	224.5	2.2	249	6	US-10-793-626-2910 Sequence 2910, App
30	224.5	2.2	341	6	US-10-793-626-3202 Sequence 3202, App
31	221	2.2	2261	6	US-10-995-561-600 Sequence 600, App
32	220.5	2.2	266	6	US-10-467-657-1150 Sequence 1150, App
33	213	2.1	265	6	US-10-793-626-2500 Sequence 2500, App
34	210	2.1	360	7	US-11-082-389-90 Sequence 90, App1
35	210	2.1	360	7	US-11-082-389-92 Sequence 92, App1
36	209.5	2.1	257	7	US-11-074-176-220 Sequence 220, App
37	209	2.1	265	6	US-10-793-626-1976 Sequence 1976, App
38	206	2.1	465	7	US-11-082-389-284 Sequence 284, App
39	205.5	2.1	577	7	US-11-082-389-134 Sequence 134, App
40	204	2.0	358	6	US-10-467-657-3446 Sequence 3446, App
41	202	2.0	245	6	US-10-467-657-2686 Sequence 2686, App
42	202	2.0	365	7	US-11-082-389-132 Sequence 132, App
43	202	2.0	376	7	US-11-082-389-122 Sequence 122, App
44	202	2.0	376	7	US-11-082-389-124 Sequence 124, App
45	202	2.0	479	7	US-11-082-389-130 Sequence 130, App

ALIGNMENTS

RESULT 1
US-10-995-561-968
Sequence 968, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OR INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OR INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 968
LENGTH: 1023
TYPE: PR1
ORGANISM: Homo sapiens
US-10-995-561-968

Query Match 24.9%; Score 2490; DB 6; Length 1023;
Best Local Similarity 48.9%; Pred. No. 2.5e-174;
Matches 510; Conservative 194; Mismatches 313; Indels 26; Gaps 9;

QY	495	MKSKDNRIKLMNELLINGIKYLAWEFLAKDKYLAIRQELKYLKKSAYLSAVGTFTWV	554
DB	1	MROKDSRRLRTSSILRNSKTIKFKHGWGAFDLRVGLRGDELALRTSGILFVSIVSQ	60
QY	555	CTPEFLVALCTFAVYVITIDENNIIIDAOFAFVSLAFNLIRPMLIPWISSIYOASVSLK	614
DB	61	VSTFLVALVFAVHTTVAB-NANNAECAFVTLVNLINKQAFLPFSIHSIVQARSPD	119
QY	615	RLRIFLSHEELPDSIERRPVKDGGSNTSYVNAETFWASDPPTLNGITFSPBGAIV	674
DB	120	RLVTFLELREVDGVDDSSSSGSAAGDCCTTISATFWQESBPCLHRLNLTVPQCLL	179
QY	675	AVVGQVCGGKSSLSALLAEMDYEGHVAIKGSVAVYPOQAMTQNDLSRENILFGOLEE	734
DB	180	AVVGPOVAGKSSLSLSLIGELSTVEGVSLEGAAVYPOEAMVQNTSVENVCFGQELDP	239
QY	735	PYRSVYAOACALPDEILPSGDRTEIGEGVNLSSGQORVSLARAVYGNADITLPDDP	794
DB	240	PMERVLAKALPDPDVSFPEGHTISIGEGNNLSGGQKRLSLAARAVYKKAIVYLDDP	299
QY	795	LSAVDAHVGHIFENYIGPKMLKNKTRILVTHSMSTLPQVDVITYVSGKISEMSGYOE	854
DB	300	LAALDAHVGHVNOVYIGPGILQGTTRILVTHALHLLPQADWITYVLNCAIEMSGSYOE	359

```

Qy 855 LLARDGAFAEFLRTYASTEOQDAENGVTGSGPGKEAKOMENGMVLTDSAGK-OQOR 912
Db 360 LLOKRGKLVCL-----DQARQPGDRG-BEGTEPGSTKDPRG-----TSAGRRPELRR 407
Qy 913 QLSSSSYSGDISHNHNSSTAELOKAEAKEET-----WKLMEADKAKQNGOVLSYVWMDMK 968
Db 408 ERSLKS-----VPEKDRITSEAO--TEVPLDDPPDAKMPAGK-DSIQGKVATVHALYLR 460
Qy 969 AIGLFISFLIFLPMCNHVSALASNYMLSLMTDDPIVNGTOEHKTVLSVYGALISQGI 1028
Db 461 AVGPPLCLYALFLPLCQGVASFQCGVWLSWADDPAGVGQGTQALNRGIFGLGCAI 520
Qy 1029 AVGEYSNAVSIGTILASRCHVDLHLSLRSPPGFEPTSGCNTVNRFSKELTVDSMT 1088
Db 521 GLFASMAAVLVLGARASRLFLQRLMDVRSPISEFFERTPIGHILNRFSSKETDVVDIDP 580
Qy 1089 EVIMFGSLFENVGACIVILLATPIAIIIPGLIYFFQRFVYVASSRDLKLESVR 1148
Db 581 DKLSLIMYAGLIEVSLVAVATPLATVALPLFLYAGFQSLYVSSCOLRLESASY 640
Qy 1149 SPVYSHNETLLGVSVIRAFEEORFIHQSDLKVDENOKAYPSIVANRWLAVLECVN 1208
Db 641 SSVCSHMAETFGSTTVARAFRTQAPFVAGNNARVDESQRISFPRLVADRWLAATVELLN 700
Qy 1209 CIVLPALFVITSHHSISAGIVGSYSTLOYTTYLMLVMSSEMETNIYAVERLKEYS 1268
Db 701 GLVPAATCAVLSKHAISAGLVGFSVSAALQVOTQLQVWVRNMTDLSENSIYSVERMODYA 760
Qy 1269 ETEKEAPMOIOETPSPSPWQVGRVPEFNNYCLARYBEDFVLRIHNTVINGEKVGIYGR 1328
Db 761 WTPKEAPWRLPTCAAPRPWPGQGFEBDFGLRTRPELPLAVQGVSEFTHAGEKVGIVGR 820
Qy 1329 TGAKSSLLTGLFRINESAGEIITIDGINAKIGLHDLRFKITIIPQDPVLFSSGLRML 1388
Db 821 TGAKSSLASLALQEAEBGIMWDGPIAHVGHTRISISIIPOPIIFPSSLRML 880
Qy 1389 DPFQSYDEEYVWLSLELAHKDFVSALPDKLDHCAEGENLVSQOROVLCALRLKRT 1448
Db 881 DLLOEHDDEALWMALETVQLKALVASLPGQLQYKCADGEBLSVGOKQLCLARALRK 940
Qy 1449 KIIVLDATAVDELTDDLOSTIRTOPEDECTVLTIAHMLTINDYTRVYVLDGEIOEY 1508
Db 941 QILILDEYTAADVGTETLQWQAMLGSWFACTVLIHRLRSVMDCAVAVMDKQVABES 1000
Qy 1509 GAPSDDLQORGLFYSMADAGLV 1531
Db 1001 GSPAQQLAQKGLFYRLAQESGLV 1023

```

```

RESULT 2
US-11-090-439-24
; Sequence 24, Application US/11090439
; Publication No. US20050266442A1
; GENERAL INFORMATION:
; APPLICANT: Squillace, Rachel
; APPLICANT: Weiner, Michael P.
; TITLE OF INVENTION: Immortalized Human Tuberos Sclerosis Null
; TITLE OF INVENTION: Angiomyolipoma Cell and Method of Use Thereof
; FILE REFERENCE: 24318-502
; CURRENT APPLICATION NUMBER: US/11/090,439
; PRIOR FILING DATE: 2005-03-25
; PRIOR APPLICATION NUMBER: 60/556,344
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 1581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-090-439-24

```

```

Query Match 21.1%; Score 2115; DB 7; Length 1581;
Best Local Similarity 31.9%; Pred. No. 1,4e-146;

```

```

Matches 527; Conservative 305; Mismatches 564; Indels 256; Gaps 37;
Qy 32 CFQNTVLVWVPCFYLMACFPFPFL-YLSRDRGYIQMTPLNKTKTALGF-----LMLVCM 86
Db 26 CFVDALVWVPHVFLFTFPFLFGWGSQSSKVIHNS-----TWLHFPCHNLRLMLTF 79
Qy 87 ADLFYSFWESRGIFFLAPV-----FLVSP-----LIGITTLATFLIQLERRGVSSG 136
Db 80 MLFLVLCETAEAGILSDGVESHHLIYMPAGMAFMAVSVV--YYHNITETSNPKLLI 137
Qy 137 IMLFVILVAVCALAIRSKIMTALKEDAQVDLFRDITF-----VYFSLILQVLVS 189
Db 138 ALLVYMTLAFI-----TKIKFVKFLDHAIGFSQPLCTGLLVLYLGMKLIVE--VN 188
Qy 190 CFSRSPLEFSETIHPNCPES-----SASFSLRTFWMITGLIVRGYROPLEG 238
Db 189 VIRARAIIFKTRPEVAP--PEDJODLVRFLOPVLNLSKTYMMWNAFIKAKKPI-- 245
Qy 239 SDMLSNKEPTSEOVVULYK--NMKKECAKTRKQVVKVYSSKDPAPQKSSKVDANE 295
Db 246 -DLRAIK-----LPIAMRALTNYORLC-----E 268
Qy 296 EVELIYKSPQ--KEMNPSLFPVLYKTEGPFYFMSFPKAIHDLMMFSGPOLIKLIRV- 353
Db 269 AFDQVQRKDIOGTQGARAIWQALSHAFGRVLVSTRILADLGFAGPLCIFGIVDHIG 328
Qy 354 --NDTKAPDWQ-----GYFYVLLFVTACIQTVLVHOYFHICFVSGRIK 396
Db 329 KENDVPQPKQFLGVYFVSSGEFLANAYVLAVLLFLMLLQRTFLOQSYVAILEGINLR 388
Qy 397 TAVIGAYYRKA--LVITNSARKSSTVEIYNLMSVDQREFMDLATTYNNIMSAPOYLIA 454
Db 389 GAIQTKLYNKMHLSTNLSGEMTACQICNLVADITNOIMFPFLCPNLMAMVQIIVG 448
Qy 455 LYLMLNGSVLAGVAVMLVMPVNAVMAMKTTQVQVAMKSKDNRIKLNELINGIKV 514
Db 449 VILLYVILGVSAALGAIVITLILAPVOYFVAKTUSQAPRSTLEYSNERIKQTNEMLRGKL 508
Qy 515 LKLYAMELAFQDKVLAIRQEBLKVKLSAYLSAVGTFTWVCTPPLVALCTPAVVYTTDEN 574
Db 509 LKLYAMENIRTRRETRRREKEMTSLRAPALYTSISIMNTAIPAAVLTITVGVSPFKE 568
Qy 575 NILDQAQVYSLAFNLIRPLNLIPWISSIYQASVSLKRLIFLSHELEPDSI--- 630
Db 569 ADFSPVAFASLSLPHILVPLFLSSVSVSTVALVSQGLSSEFLSABIREBQCAPHE 628
Qy 631 -----ERRPVK-----DGGGN--SITVANATPTW 653
Db 629 PTPQPAKYQAVPLRVNRRRPARBDCRGLTGFLQSLVPSADGDADNCCVQIMGYFT 688
Qy 654 ARSDPTLNGITFSIPGALVAVVGVGCGKSSLSLALAEAMDVEG----- 700
Db 689 TPDSIPTLSNTTIIIPRQGLTMIVGVGCGKSSLLALGMMQVSGAVFWSLSLPDSRIG 748
Qy 701 -----HVAIKGSVAVVPPQAMIQNDLSRENILFGQLEBPYKSVIQAALL 747
Db 749 EDPSERETATDLDIRKGPVAYVQKPPMLNATVENIIFESPFRKORYGVATEACSLQ 808
Qy 748 PDLEILSGDRTEIGEAGVNLSSGQKORVSLARAVYNSADLYLPDDPLSANDAVKGIIF 807
Db 809 PDIDILPHGOTOIGERGINISGQORISARALYOHANAVPLDDPESALDHLSDHLM 868
Qy 808 ENVIGPKGMLK-----NKTRLIVTHSASYLPQVVIIVMSGKISEMGSYOELARDGAF 862
Db 869 Q-----AGILELRLADDKRTVVLVTHKQYLPHADWITIAMKDGTLQREBGLKDPQSEQL 923
Qy 863 AEFRTYASTEOQDAENGVTGVSFGPEAKQOMENGLVYDSAGKLOQOLSSSSSYSG 922
Db 924 FEHWKTLMN--RQDELKEKETY-----ERKATEPPOG-----LSRAMSSRDG-- 965
Qy 923 DISHNHNSSTAELOKAEAKEETWKLMEADKA-----QTGVKLSVYMDYMAIGLIFISFL 977
Db 966 -----LQDEEEEBEBAESEBDNLSMHLQRAEIPWRACAKYLSAGILLLSL 1014

```

```

OY 978 SIFLPMCHNVSALSNWMLSLMTDDPIV-----NCTOHT-----KYRLSYGMLG 1023
Db 1015 LVFSOLLKMWLVALDIDYWLAKTIBDALTTLTPAANCSSJSECTLLDQIVYAWFVYLSG 1074
OY 1024 ISOGIAVEGYSMAVSIIGGILASRCLHVDLLSHLSLPSMPSFFERTPSGNLVNRFSEKLDTV 1083
Db 1075 I---VLGVTSVTVIEWTGTAKVAKRLHRSLSLNRILLAPRFFETPLGSLNRPSSDCNTI 1111
OY 1084 DSMPEVYIKMFMSGLFNVIACIVYLLATPLIAITIIPLGLIYFVQRFVYASSRQKRL 1143
Db 1132 DQHPSTLECTSRSLTLCVSLAAVTSYVTPVFLVALLEPLAIVCYFIQKXPFVASRDLQOL 1191
OY 1144 ESVSRSPYXSHNETLLGVSIVARFEBOERFHHOSDKVDENOKXYPSIVANRMLAVRL 1203
Db 1132 DDTYOLPLSHFAETVEBGLTTIRAPRYEARFOQKLETTDSNNNLASFLTPAANRLEVRM 1251
OY 1204 ECVGNCIVLFAALFAVIS--RHSLSAGLVGLSVSYSLQVTTYLVNMLVMSSEMETNIYAV 1261
Db 1252 EYIGACVVLIAAVTSISNSLHRELSAGLVGGLTVYALVNSYLVMMVRNLDLMELQIGAV 1311
OY 1262 ERKXEYSTEKE-----APMOIGETAPSPSPQVGRVPEFRNYCLIRREDDFLYLRHNV 1315
Db 1312 KRHHGLMTKEESYEGLLAPSLI-----PKWPPDQKIQIONLTVRXYDSSLKPYLKHVNA 1366
OY 1316 TINGEKYGIQORTAGKSSLLTGLFRINESAEGEIIIDGINIAKTGLHDLRPFKTTIPQ 1375
Db 1367 LISPOKIGICGRTOSGKSSSLAFAFRVNDTREGHIIIDGIDIDIAKPLHTIRSRSLILO 1426
OY 1376 DPEVFSGLRNMLDPEQYSDBEVWTSLELAHLKDQFVASLPDKLDHECABEGENLVSQOR 1435
Db 1427 DPEVFSGTIRFNLDERKCSOSTLWEALEIQKLKVLVYALPGGDAIITEGGENPSSQOR 1486
OY 1426 QLVCLARALKKTKLVLDERTAAVDETDLIGSTIRTOEDCTVLTIRLNTINDYT 1495
Db 1487 QLFCLARAFVKRTSLIFIMDEATASIDMAITEMILOQVWMTAADRIVTVIARVHTILSAD 1546
OY 1496 RVIYLDKGEIOEYCAPSDILOOR--GLPFSMAK 1526
Db 1547 LVIVYKRAILIEFDKEPTEKLSRKOSVFASFVR 1578

RESULT 3
US-10-995-561-527
: Sequence 527, Application US/10995561
: Publication No. US20050272054v1
: GENERAL INFORMATION:
: APPLICANT: CARGILL, Michele et al.
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
: TITLE OF INVENTION: DETECTION AND USES THEREOF
: FILE REFERENCE: CU001559
: CURRENT APPLICATION NUMBER: US/10/995.561
: CURRENT FILING DATE: 2004-11-24
: NUMBER OF SEQ ID NOS: 85702
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 527
: LENGTH: 1588
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-995-561-527

```

	Query Match	21.0%	Score 2101.5	DB 6	Length 1588
	Best Local Similarity	31.8%	Pred. No. 1.4e-145		
	Matches	527	Conservative 305	Mismatches 564	Indels 263; Gaps 39;
OY	32	CEQNTVWVCFTYMACEPFYFL-YLSRHGRGYQMPLNKTALGF---	LIMVYC	86	
		:::::		:	
Dd	26	CFVDLNVPHVFELTFTFPILEFGSSGSSKVVHHS-----TWLFPGHNIRMTLT		79	
		:::::		:	
OY	87	ADLPYSFMERSRGTFPLAVY-----EIVSPT-----LIGITTLATFLIQERRRQVSGS		136	
		:::::		:	
Dd	80	MLPFLVCEIIEGLISDGVESHHLALTPMGAMFMAAVTSVV--YYNIITSNFPKLLI		137	

QY	137	MLFMFWALVACALALRSKMTMTLKBDAQDLPEDITF-----YYFSLLLIQVLVS	189
Db	138	ALTYWYTLAFI-----TKTJFKVKFLDHAIGFSQLEFCLTGLLVITIGMLLVE--VN	188
QY	190	CFSPSRPLFSSTIHDPNPCPS-----SASFSLRITFWMTGLIYRGYQOPLEG	238
Db	189	VIRVRRYIFFTPREVKP-PEDLODLGRFLQPPVNLISKGTWMMNMFITAHKKP1--	245
QY	239	SDLSLKNEDTSEOVVPVLVK--NMKCEAKTKQPVKVYVSSKDAQPKRESSKVDANE	295
Db	246	-DEAIGK-----LPIAMRALINYORLC-----E	268
QY	296	EVEALIVYSPQ-KEMNPDLFVKLYKTRQPYPLMSEFFPKAHDLMWFSGPOLLKLIKPV-	353
Db	269	AFDQVQRKDIOGTQGAIRIOWALSHARGRLVISTRIIADLDIGFAPLCLIFGVHLG	328
QY	354	--NDTKAPDWQ-----GYFYTVLLFVTAQTLQTVLHQFYHIFVFSGRMK	396
Db	329	KENDVPQKTFQGLGVYFVFSQEFANAYVLALFLFALLLQRTFLOASYYAIEFGIMLR	388
QY	397	TAVIGAVYRKA--LVITNSAKKSTGEIVNLMSVDQRPMDLATYINMISAPLOVLA	454
Db	389	GAITOTKYNKTMHLSTSLNSMGEMTAQICNLVALIDTQLMWFFPLCPNLMAMPQIIIG	448
QY	455	LYLDMNLGPESVLAGVVMVLMVNVNMMAKTKTYQVAHMKSKDNRIKLINELINGIV	514
Db	449	VILLYILIGVSALIGAANIIILAAVOYFVAFKLSQAQSRSTLEYSNERKQNMELRGIKL	508
QY	515	LKLYAMBLAFKOKYLAIROEBLKVLKKSAYISAVGTFPTWCTPPLVALCTPAVYVTTDEN	574
Db	509	LKLYAMENIEFTRVETTRRKEMTSLRAFAITYSISIFENNTAIPIAAVLITFVGHSFPKE	568
QY	575	NILDAQAFVSLAFENLIRPELNTLPMWISIVQASYSIKRLAFLSHEELBPSI----	630
Db	569	ADSPBSVAFASLSLPHLLVTPFLILSSVNSITVALVSVQKLSLSELSAEIRBEOCAPE	628
QY	631	-----ERRPVK-----DGGGTV-SITVANAPETW	653
Db	629	PTPGPASKYQAVPLRVNRRRPAREDRGTLGTPLOSLVPSADGDADNCCVINGGYETW	688
QY	654	ARSDPPLNGTTFEIPEGALVAVVGVGCGGSSLSLALAMDMQVEG-----	700
Db	689	TPDGIPTLSNITTIRIPGQLTMIYGVCGCGSSLLLALGMOQYSGAVFSSLPDSBIG	748
QY	701	-----HVALKGSVAAYVPOQAMIONDSLRENILFEGCOLSEBYRYSVIOACALL	747
Db	749	EDSPREBETADDIRKRGPAVAYSOQRPWLMAVEENIIFESPKNQRYCMVIEACSLQ	808
QY	748	PDLFILPSGDRTEIGEGKGNVLSGGOKORVSLARAVSNAIDYLFDDPLSAVDAHVKHIF	807
Db	809	PDLILFPHGDOQTQIGERGIMLSGGORORISVABALYOHANVFLDDPSALDIHLSHLM	868
QY	808	ENVIGRPGMK-----NKTRLIYHMSYLPQVUVIIVMSGCKISEMSGYELLARQAF	862
Db	869	Q-----AGILELRLDDKRIYVVLVHKKIQTYPHAWIITAMKQGTIQREBTLDFQSRSEQL	923
QY	863	AELFRTASTEOQDAEENGVTGVSGPGEAKQKQENGMLYTDSAGKOLQORLSSSSSYSG	922
Db	924	FEHKHTLMN-RQDELKEITY-----ERKATEPQG-----LSRAMSBDGL--	965
QY	923	DISHHNSSTAELOCAEAKKEETWKLMAZKA-----OTGOVKLSVWDYMKATGLFISFL	977
Db	966	-----LOOEHEEEBBAASEEDNLSMILHRAELPWRACAKYUJSSAIGILLST	1014
QY	978	SIFLFMCNHNVSALANSYMLSLMTDDPIV-----NGTOEHT-----KURLSVYGAIG	1023
Db	1015	LVFSQGLKHVYVAIDVWIAKMTDSALTLPRAANCGLSQECLTDQIVYVAVFVVLCSIG	1074
QY	1024	ISGQIAVFGISMAVSGIGILASRCLAYLDLHSILRSPMSFFERTPSCGLVNRPSKELDTV	1083
Db	1075	I-----VLCLIVSVTEWGLKVARKLHSLNRIITLAPMRFEETPLOSISINFSSDCNTI	1131

966	-----LQBBEEBEEBSEEDDNLKMLHOPAEIPWRAKATYSSAGILLLSL	1014
Qy	978 SIFLEMCHVSALASNYWLSLMTDDPIV-----NGTOEHT-----KVRLSVYGALG	1023
Db	1015 LVFSQLKHKHVLVAIDYWLAKMTDSALTLPRAARNGSLSGECTLDQIVYAVWVPVLSIG	1074
Qy	1024 ISGSIANFGVSMASVIGSILASRCLHYDLHLISLRSPMSFFERTPSGNLVNRSEKELDTV	1083
Db	1075 I--VLCVLTSVYEWGTLKVAKKLHSLNRIIILAPMRFEETPLQSLINRFSSDNTI	1131

QY 1084 DSMPEVVKMFMSGLFNVIAGACIVILLATPIAIIIPPLGIYFFVQRFYVASSRQKRL 1143
DB 1132 DQHPSTLECLSRSTLLCVSALAVISYVTPVPLVALPLAIVCYFIQKYFVASSRDLQOL 1191
QY 1144 ESVSRSPVYSHFNLTIGVSVYIRAFEBQERFIHOSDKVDENQAAVYSIVANMMLAVRL 1203
DB 1192 DDTQLPLSHFAETVEGLTITIRAFREARFOQKLEBTDSSNNIASLFLITANMMLVYRM 1251
QY 1204 -----ECVNCJVLFPALFAVIS--RHSLSAGLVGSVSYSLQVTTYLMVLVMSSEM 1254
DB 1252 ATPLPQEVYIGACVCLAAVAVISNSLHRELASAGLVGLTYALMVSVYLMVWRNLADM 1311
QY 1255 ETNIVAVERLKEYETEKE-----APMOIQETAPSSWPOGVREFRNYCLARREDIDF 1308
DB 1312 ELQGVAVRHHGLKTESEYEGGLAPSLI-----PKMWPOQGIQIQLNLSVRYDSLSKP 1366
QY 1309 VLRIHNTINGEKGVYVGRGAGKSSLTGLFRINSEABEGLIIDGINIKIGLHDIRF 1368
DB 1367 VLKVNMLISPGQKIGICGRGSGKSPSLAFPRVWDFEGHIIIDGIDAKLPLHTLRS 1426
QY 1369 KITTIPODPVLFSSSLMNLDPFQOYSDDEVYTSLELAHKDFVSALPDKLDHCEABEGE 1428
DB 1427 RLSTIILDOPVLFSTIRHNDPERKCSSTIMBALIQLTVYKALPGJDAIITBEGE 1486
QY 1429 NLSVGRQVCLAPALKRKIKILVDEATAVDETDLIOSTIRTOPECTVLTIAHRL 1488
DB 1487 NFSOGQOLFCLARAFVAKTISIFIMDEATASIDMATENIIQKVWTFADRTVVATIAHRY 1546
QY 1489 NTIMDYTRVYLDKGEIOEVGAPSDILOQR--GLFYSMAK 1526
DB 1547 HTILSADLVIVLKRGALEFDPKPEKLSRKDSVFASFVR 1585

RESULT 4

US-11-090-439-26
Sequence 26, Application US/11090439
Publication No. US20050266442A1
GENERAL INFORMATION:
APPLICANT: Squillace, Rachel
APPLICANT: Weiner, Michael P.
TITLE OF INVENTION: Immortalized Human Tuberos Sclerosis Null
TITLE OF INVENTION: Angiomyolipoma Cell and Method of Use Thereof
FILE REFERENCES: 24318-502
CURRENT APPLICATION NUMBER: US/11/090,439
CURRENT FILING DATE: 2005-03-25
PRIOR APPLICATION NUMBER: 60/556,344
PRIOR FILING DATE: 2004-03-25
NUMBER OF SEQ ID NOS: 62
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 26
LENGTH: 1581
TYPE: PRN
ORGANISM: Homo sapiens
US-11-090-439-26

Query Match 21.0%; Score 2100; DB 7; Length 1581;
Best Local Similarity 31.7%; Pred. No. 1.7e-14;
Matches 524; Conservative 306; Mismatches 567; Indels 254; Gaps 37;

QY 32 CFQNTVAVWVCFYLMACFPFYFL-YLSRHDRGYIQMTPLNKTALGCF-----LMIYCW 86
DB 26 CFVVALNVVPHVFLIFITPILFIIGMGSSQSKVHIHS-----TWLHFPCHNLRMLITLF 79
QY 87 ADLFYSFMRSGRGIFLAV-----FLVSPR-----LGIITLLATNELLQLERRKGVQSSG 136
DB 80 MLPLVAVCEIABGLISDGVSSHLLHLYMPAGAMFAAVTSV--YNNIETSNFPKLIT 137
QY 137 IMLEFWALVACALALSKIMTALKEDAOVLDFFDITF-----YVYFSLLIQVLVS 189
DB 138 ALVYVWTLAFL-----TKTIKPVKLDHAIIGFQLFCULGLVILYKMLLVE--VN 188
QY 190 CFSRSPVYSHFNLTIGVSVYIRAFEBQERFIHOSDKVDENQAAVYSIVANMMLAVRL 1204
DB 193 DDTQLPLSHFAETVEGLTITIRAFREARFOQKLEBTDSSNNIASLFLITANMMLVYRM 1252

DB 189 VIRVRRIYFFKTPREVKRPPEDLODLGVRFLOPFVNLPESKGTYYMMANFITAHHKPI--- 245
QY 240 DLMSLNKEDTSEQVVPVLVK---NMKECAKTRQOPVKVYSSKDPAPQESSKVDANEE 296
DB 246 DLRAIGK-----LPIVMBALNTNYRLC----- 269
QY 297 VEALIVKSPQ--KEMNBSLFKYLKTFGRPYFLMSFFPAIHDLMFSGPQIIKLIRY-- 353
DB 270 FDAQVRKDIOGTQGARAIWQALSHAFGRRLVLSSTFPIILDLLGFAGPLCIFGVHDIGK 329
QY 354 -NDTKAPDMQ-----GYFYTLFVTACLOTLYLHQGFHICFVSGMBIKT 397
DB 330 ENDFQPKTQFLGVYFVSSQBFANAVYLALELALLORTFLQASVYAIETGINDRG 389
QY 398 AVIGAVYRKA--LVITNSARKSSTVGEIIVNLSVDAORFMDLATYIMNIMSAPLQVILAL 455
DB 390 AIQFKIKNKLIMHLSLNSMGEMTAGICNLVADTNQMLWMPFLCENLAMPQIIIVGV 449
QY 456 YLLMLNIGPSYLAGVAVMTLVPNVAVAMKTKTYQVAHMKSKONRIKLANEILNGIKYL 515
DB 450 ILLYYIIGVSAIIGAAVYIILAPVQYFVATKLSQAQRSTLEYSNERLQKTEMLRGIKLL 509
QY 516 KLYAMELAFDKVLAIRQBELKVLKKSAYLSAVGTFTWVCPPVLVALCTFAVYVTTIDENN 575
DB 510 KLYAMENIFKTRVETTRKKEWTSIRAPAIYTSIFNNYTAIPIAVALITFGHVSFFKEA 569
QY 576 ILDAQTAFVSIATFNILRPPLNLIIPWISSIVQASVSLKRLRPLSHELEPDSI----- 630
DB 570 DFSVSAFASISLPHILVTLPLFLSSVVRSTVAVLVQKLSBFLSABEIREQCAPHEP 629
QY 631 -----ERRPVK-----DGGGTN--SITVRNATPTWA 654
DB 630 TPQGPASKYQAVPLRVNVRKRPAEEDCRGLTGPLQSLVPSADGADCCVQIMGCVFTWT 689
QY 655 RSDPPTNGITFESIPAGALVAVVGVGCGKSSLSLSALAEADKDEG----- 700
DB 690 PDGIPLSNITIRIPRQLTWIVGQVCGKSSLSLALAGEOKVSAVMPSSLPDSB1IGE 749
QY 701 -----HVAIKGSVAVVPOQAMIONDSIRENILFSCOLEBPPYRSVIOACALLP 748
DB 750 DSPERETADTDLDRKRGPAVYASQKPMNLNATVENIIPESPNNKRYKNVITACSLQIP 809
QY 749 DLEILPSGDRTEIGEEKVNISGGQKQKRVSLARAVYSNADYILFPDPLSAVDAYHKKHFE 808
DB 810 DIDILPHGDOQTOIGERGINISGGQKQKRVSLARAVYSNADYILFPDPLSAVDAYHKKHFE 808
QY 809 NVIGPKKMKL-----NTRILIVYHSMGYLPQVUVIYIMSGKSIEMSGYOBLLARDGAF 863
DB 870 -----AGILELRDRDKRTVAVLVTHKQLLPHADWIIAMKDGSTIQRBGLKDPORSECOLF 924
QY 864 EFLRTVASTEOQDAENGVTVGVSGPGEAKQKEMNGMLVTDSAGKOLQROLSSSSSYSGD 923
DB 925 EHKMTILMN-RQDOQLKEHYI-----ERKATEPQG-----LSRAMSSSDGL--- 965
QY 924 ISRRHNSIABLOKAEKKEETWKLMEADKA-----QTGQVKLASYWDMYKALGFLISFLS 978
DB 966 -----LQDEEBEEREEAEBEEDNLSMHLQBAEIPWRACAYLSAGGILLISL 1015
QY 979 IELFMCHVSAIANNVLSLMTDDPIV-----NTOGEHT-----KRLSYGALGI 1024
DB 1016 VFSQLKRMVYLVADLYMLAKWTDSALTITPAARNCSLSQECTDDQTYAVAMFYVLCISGI 1075
QY 1025 SQGIADVGSMAVSIIGGILASRCHADVLLSHLSIRSPMSFFERTPSGMLVNVFSKELDTVD 1084
DB 1076 ---VLCVTSVTVWGTGKAKAKRLHRSLLNRILIAIPMRFPETTPGLGSLNRFSSDCNTID 1132
QY 1085 SMPEVVKMFMSGLFNVIAGACIVILLATPIAIIIPPLGIYFFVQRFYVASSRQKRL 1144
DB 1133 QHISTLECLSRSTLLCVSALAVISYVTPVPLVALPLAIVCYFIQKYFVASSRDLQOL 1192
QY 1145 SVSRSPVYSHFNLTIGVSVYIRAFEBQERFIHOSDKVDENQAAVYSIVANMMLAVRL 1204
DB 1193 DDTQLPLSHFAETVEGLTITIRAFREARFOQKLEBTDSSNNIASLFLITANMMLVYRM 1252

Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C0001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 969
; LENGTH: 410
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-969

Query Match 9.7%; Score 969; DB 6; Length 410;
Best Local Similarity 50.8%; Pred. No. 1.4e-63;
Matches 197; Conservative 66; Mismatches 121; Indels 4; Gaps 2;

QY 495 MKSKDNRIKLMNEITNGIKVILYAMELAFKDKVLAIRQEBELKYLKKSAYLSAVGTPTWV 554
DB 1 MRQKDSARLRITSLRNSKTIKFGHMGCAFIDRVGLRGQELGALRTSGLLFSVLSVFG 60
QY 555 CTPLVALCTFAVVTITDENNILDQTAFAVSLAFNLIFFPLNLPWVSISSIVQASVSLK 614
DB 61 VSTFLVALVFAVHTLVAE-NANNAEKAFVTLVNLINKAQAFLPFIHSLQVARSFD 119
QY 615 RLRLFLSHEELPDSIERRPVYKGGGNTSTVRNATFWARSDPTLNGITFSIPEGALV 674
DB 120 RLVLFLCLIEVDPCGVDSSSGSAAGKDCITQSTAFMSQESPCLHRIMILVPOGCLL 179
QY 675 AVGVGVGGKSSLSALLAEMDKVEGVAIKGVAVYVPOQAMIONDSRENILLEGCOLEE 734
DB 180 AVGVGVGGKSSLSALLAEMDKVEGVAIKGVAVYVPOQAMIONDSRENILLEGCOLEE 734
QY 735 PYRVSIVQACALLPDLBILPISGDRTEIGEKVNLSSGQKORVSLARAVYSNADIVLPDP 794
DB 240 PWLERVLKCALQPDVDSFPEGHTISIEGQGMNLSGGKQSLSLARAVYRRAAVYLLDDP 299
QY 795 LSAVDAYGKAIFFENYVGPCKMLKNTIILVTHSMSTYLPQVDVYIWSGKISMGSTOE 854
DB 300 LAAIDAYGQHVFQVIGPGILQGTTRILVTHALHILPQADWIIIVLANGAIAMGSIYOE 359
QY 855 LLAADGAFAEFLRTYASTEOEDAEENG 882
DB 360 LLQKKGALVCLL---DQARQPDGRGEGG 384

RESULT 7
US-10-613-744-14
; Sequence 14, Application US/10613744
; Publication No. US20050272093A1
; GENERAL INFORMATION:
; APPLICANT: Mackinnon, Roderick
; APPLICANT: The Rockefeller University
; TITLE OF INVENTION: Assays for Screening Compounds Which Interact With
; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
; TITLE OF INVENTION: Channel Proteins, and Uses Thereof
; FILE REFERENCE: 018512-002901US
; CURRENT APPLICATION NUMBER: US/10/613,744
; CURRENT FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US/09/275,252
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: US 09/045,529
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/054,347
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: WO PCT/US99/06307
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 14
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-613-744-14

Query Match 8.8%; Score 881.5; DB 6; Length 407;
Best Local Similarity 46.1%; Pred. No. 3.6e-57;
Matches 167; Conservative 75; Mismatches 95; Indels 25; Gaps 4;

QY 1551 LTRLAEPDAPRY----RTERRARFVSKGNCVAKHNRQGRFLQDVTTTLVLDLKW 1606
DB 16 LPKQARDL--PRHISRDRTKRIQRYVKKQKCNVHGNRETRYRLDITFTLVLDLKW 74
QY 1607 HTLLIFTFMFLCSWLLFAMVWMLIAFAHGLDYAYMERGITDLAPYDVPDPAEGGTNP 1666
DB 75 FNLLIFPMVYTVMLFFGMIWMLIAYIRGDM-----DHIEDSPWTP 115
QY 1667 CVTSHSPSSAFPSIIVQVTIGFGGMVTEBCPLAILLVQNIQGLMNAIMAGCTPM 1726
DB 116 CVTNLNGFVSAFLPSIETETLIGYGVYITDKCEGIIILLIQGLSIVAFVWGMFV 175
QY 1727 KTAQAHRAETLIRSKNAVITLRHGRCLCFMLRVGDLKSMISTIMQVVRKTTSPGE 1786
DB 176 KISQPKRAETLVFSTAVISMARDKCLMFRVGDILKNSHIVERAKLIKSKQTSGE 235
QY 1787 VVPLHQVDIPMENGVNGGIFLVAPLIIYHVIDSNSPLYDIAPBDLHHODLEIIVLEG 1846
DB 236 FIPNLQTDINVGYYTGDRLEFLVSPLIISHINQSSFWELSKQL-PKELEIVLILEG 294
QY 1847 VVETGTTTQARTSYLADBLIMGQRPVPIVAEEDGRYSVDYSKRGNTIKVETPLCTARQL 1906
DB 295 MVEATGMTCQARSYSYIISEILMGYRFPVLLLEDGFYEVDSNPHETVETSPLSAKEL 354
QY 1907 DE 1908
DB 355 AE 356

RESULT 8
US-10-995-561-967
; Sequence 967, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; APPLICANT: The Rockefeller University
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C0001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 967
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-967

Query Match 8.0%; Score 799; DB 6; Length 268;
Best Local Similarity 56.3%; Pred. No. 2.1e-51;
Matches 151; Conservative 56; Mismatches 61; Indels 0; Gaps 0;

QY 1264 LKESSETEKAPMGIQETAPSSMPOYGRVFRFVYCLRYREDLDFVLRHINVTINGGKV 1323
DB 1 MODYAMTPKEXPWRLPTCAQAPMPQGGQLBFRFGRLRPBPLAVQGVSFKTHAGSKV 60
QY 1324 GIVRTAGKSSLTGLFRINESAGEIILIDGIMIAKIGLHDLRFKTIIPQDVLFPSSG 1383
DB 61 GIVRTAGKSSLASGLRLQEALEGWIDGVYIAVGLHTLASRSIIRQDILFPSSG 120
QY 1384 LRWULDPSSQYSDSEVWTSLELAHLKDFVSALPKLHDECAEGEENISVGGROIVCLARA 1443

Db 121 LRNNLLDLSHSDRAIMAALETVOALKALVSLPGLOKYCADRGEDLSVGGKOLLCLARA 180
Qy 1444 LRRTKILVDEATAVDLETDDLOSTIRQFEDCVLTIAHRLNTIMDTRYIVLDKG 1503
Db 181 LRRTQILLIDETAADVDPGTELOQMAMLSWFOACTVLLIAHRLSVMDCARVLVMDKG 240
Qy 1504 EIOEYGAUSDLLQORGLFYSNADAGLV 1531
Db 241 QVAESGSPAQLAQKGLFYRLAQSGLV 268
RESULT 9
US-10-613-744-13
; Sequence 13, Application US/10613744
; Publication No. US20050272093A1
; GENERAL INFORMATION:
; APPLICANT: Mackinnon, Roderick
; TITLE OF INVENTION: Assays for Screening Compounds Which Interact With
; TITLE OF INVENTION: Cation Channel Proteins, Mutant Prokaryotic Cation
; TITLE OF INVENTION: Channel Proteins, and Uses Thereof
; FILE REFERENCE: 018512-002901US
; CURRENT APPLICATION NUMBER: US/10/613,744
; PRIOR FILING DATE: 2003-07-03
; PRIOR APPLICATION NUMBER: US/09/275,252
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: US 09/045,529
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 09/054,347
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: WO PCT/US99/06307
; PRIOR FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-613-744-13
Query Match 7.9%; Score 788.5; DB 6; Length 391;
Best Local Similarity 43.7%; Pred. No. 2.2e-50;
Matches 149; Conservative 77; Mismatches 92; Indels 23; Gaps 5;
Qy 1564 RTERRARAFYSKGCNCVNAKINIREQRF--QDVFITVLDKMPRTLLIFMSFLCSWL 1621
Db 35 RSRQ-RARLVSKGRCNIEFGNVAQSRFIFVDIWTVDLTKRMYKMTVFITAFIAGSWF 93
Qy 1622 LFAMVWMLIAFAHGDLYAVYMEKGITDLAPYRVDVDAEGSTNVCVTSIHSSSAFLFS 1681
Db 94 LFLGLMYVAVYAKDL-----PEFYPPDNKTPCVENINMGTSAFLFS 135
Qy 1682 IEVOVITIGFGNWTBCEPLAILILIVONIGMIMNIMLCIFMKTQAHRRAETLLFS 1741
Db 136 LETQVITIGVFRFTEGCATAIFLLIFQSLGVINSFMCAGALLAKISRKRAKTTFS 195
Qy 1742 KHAVITLIRGRCLFPMARVGLDKSMIISATIHQVAKTSSPGEVPLVHQVDIPMEGV 1801
Db 196 KNAVIRKRGKCLLIRVALRKSLIGSHIYKGLKTTITPGEETIILDQTNINEVDA 255
Qy 1802 GNGGIFLVAAPLITVYIDNSPLYLAPSDLHNODEIIVLEGVETGITTOARTSY 1861
Db 256 GNGNLFFISPLITVYIHIIDHNSPFFHMAETL--SQODELVVFLDGVVESTSATCQVRTSY 314
Qy 1862 LADIELMGRFVPIVAE-EDGRYSVDYSKFGNTIKVPTPLC 1901
Db 315 VPBEVLWGYRFVPIVSKTKGKRYVDVFNHNGKTIVETPHC 355
RESULT 10
US-10-793-626-326
; Sequence 326, Application US/10793626
; Publication No. US20050255478A1

; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: P03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 326
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-326
Query Match 5.1%; Score 514.5; DB 6; Length 548;
Best Local Similarity 27.1%; Pred. No. 4.3e-30;
Matches 157; Conservative 110; Mismatches 211; Indels 101; Gaps 16;
Qy 1002 DPIVNGTQHTKRLSVYAGLISOGIAVPGYSMAVSGIGLASRCLHVDLSILRSPM 1061
Db 12 DGVIN--NHSLNQEKFSHLGVAIGIALF-----IFLIYVPI 47
Qy 1062 SFFE---RTPSGNLVNRFSKEL-----DVTDSMIPVYK-----MFWG 1096
Db 48 EFRQYLAQWTSNKLIDIRKQLYNHLQALSVRFVANNQGVYISRYINVEQTKDILT 107
Qy 1097 SLFNVIGACIVILLAPI-----AAILIIPGL--IYFVQRFYVASSRQKLE 1144
Db 108 GLNNIMDCTITIIIASIMEFLDVKLFAIFLFPYILITVYFFGRL-----RKLTRVR 162
Qy 1145 SVSRPYSHFNETLGVSVIRAFEEQERFIHQSDLKVDNOKAYPSIVANMLAVRLE 1204
Db 163 SOLAEVQGLHERVQMSYISFAIEDNEAKNF--NNKNKPLORAFQHTRWAVSFA 219
Qy 1205 CV-----GNCIVLFAALFAVISRHSLSAGLVGLSVSYLQVTTYINMLVWMSSEMETNI 1258
Db 220 AINTVTDLGRITVYIGVSYLAT--GSITVGTLLAFVGLYELQPLRLVSSFTTLTQSF 278
Qy 1259 VAVRLKYSERKE-----APWQIOETAPSSPQVGRVFNRYCLARYEDLDVFLR 1311
Db 279 ASMDRQFQIMDEYDIDKNGIGAPRIRSK-----GQIDLKVSFFKYNENKEVLIH 328
Qy 1312 HINVTINGEKVGIIVRTGAGKSLTLGLFRINESAGEIIDIGINIAKIGLHDLRPKIT 1371
Db 329 DINTLTINKGETVAFVMSGGKSTLNLIPFYDVTOGELLIDHNVKDFLTGSLNRQIG 388
Qy 1372 IIPQDVLFSGLRNL--DEPSQYSDREVWTSLEMLKDFVSAALPDGLDHECAGE 1428
Db 389 LVQQNDILFSDTYKENIILGRF--DATDDEVVAAGMANADHPIISLNPQYDIVEGERV 446
Qy 1429 NUSVQROLVCLARALLRRTKILVDEATAVDLETDDLOSTIRQFEDCVLTIAHRL 1488
Db 447 KISGQOKRLSTARIRFLNPPVILIDEATSALDISEAIIQEALDVLSDRRTLLIAVHRL 506
Qy 1489 NTIMDTRYIVLDKGEIOEYGAUSDLLQORGLF---YSM 1524
Db 507 STTHADRIVVMENGRIVETGHQOLINRGAVEHLYSI 545
RESULT 11
US-10-467-657-1346
; Sequence 1346, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega

```

? APPLICANT: MONACI Elisabetta
? TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
? FILE REFERENCE:
?
? CURRENT APPLICATION NUMBER: US/10/467,657
? CURRENT FILING DATE: 2003-08-11
? PRIOR APPLICATION NUMBER: GB-0103424.8
? PRIOR FILING DATE: 2001-02-12
? NUMBER OF SEQ ID NOS: 9218
? SOFTWARE: SeqWin99, version 1.04
? SEQ ID NO 1346
?
? LENGTH: 622
?
? TYPE: PR1
?
? ORGANISM: Neisseria gonorrhoeae
?
? US-10-467-657-1346

```

Query Match	4.9%	Score 495.5	DB 6	Length 622
Best Local Similarity	28.7%	Pred. No. 1.3e-26		
Matches 154	Conservative 106	Mismatches 213	Indels 63	Gaps 13
Qy	1022	LGISQGIAPVFGYSNAVSIGGIIASRCKLHVDLHLSIRSPMSFFETPESGNLVNPFSELD	1081	
Dh	108	LIVIRGICRFSTYLTMTVSVWTTISKIKDMPAKMLTLSRHYHQETPBGTYLMMNLNLT	167	
Qy	1082	TVDSMIPVIMKFMGSLFNVIQACVILL---ATPLAIIIPPLGLIYFFVQRFYVASS	1137	
Dh	168	QSVSNASDIFVLTTRDTMTIVTGLTIVLLLYLNWQSLIVLWMFPLLSL---SRY---R	220	
Qy	1138	ROLRLRLEVSRSPLYVSHEN---ETLGVSVIRAPEDGERFIHQSDL-----KYDE	1184	
Dh	221	DLRKHVIDSQKSGISGTMMNVIAETHQGRVAKLFNGQHOANRPDAVNRITVRLSKITQ	280	
Qy	1185	NOKAYYP--SIVANKMLAVRLCEVGNCTIVLPALPAVISRHSLSAGLVGSYSYLQVTT	1242	
Dh	281	ATAHSPSESLIASIALA-----VIFALMQSQNGVTTIGEFMAPIVA-MLQVYA	330	
Qy	1243	YLANLVKMSSEMETNIVAVERLKEYSETEKE-----APWQIQETAPSSMPQNGRVER	1296	
Dh	331	PIKSIANISIMQMTFTFLADGVCAFLDTPPEODKCTLAPORVE-----GRISFR	379	
Qy	1297	NYCYRRERDLPFVLRHINVTINGEKKVGIORTGAGKSLTLGLPRINESAGEIITDGI	1356	
Dh	380	NVDVEYRSDGKIKALDNFEDLDRQGRKVALVGRSGGKSTIVNLLRFYEPSPGNCITDGI	439	
Qy	1357	NIAKIGLHDLPEFKITIIIPQDPVLFEFSGLRMLDPFSQYS-----DEEYVTSLELAHLKDF	1411	
Dh	440	DIADIKDCLAAQFALVQDVPFLPDTLFEVNV---RCSRPDAGAEVLSLQANLQSL	495	
Qy	1412	VSALPDKLDHECABSGENLVSGQROLVCLARALLKTKILVIDEATAVLDLTDLIQST	1471	
Dh	496	IDASPLGLHQPITGSGNSMISGGQFORVAIARAAILKDAEITLLIDEATSAIDNESEETLQQA	555	
Qy	1472	IRIQEDCTVLTIAHRLNTINDYRFVIVLDGELQEVGAPSDLLQORGLAFYSKMD	1527	
Dh	556	LERIMENRTGIIVAHRLTTVESADRIIVMDGKLTIEQTHDQLMFONG-YITMLRN	610	
RESULT 12				
US-11-082-389-148				
Sequence 148, Application US/11082389				
Publication No. US20050244935A1				
GENERAL INFORMATION:				
APPLICANT: Pompeju, Markus				
APPLICANT: Kroege, Burkhard				
APPLICANT: Schroder, Hartwig				
APPLICANT: Zeider, Oskar				
APPLICANT: Haberhauser, Gregor				
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS				
TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE				
FILE REFERENCE: BGI-113PCN				
CURRENT APPLICATION NUMBER: US/11/082,389				
CURRENT FILING DATE: 2005-03-16				
PRIOR APPLICATION NUMBER: US 09/603024				

```

PRIORITY FILING DATE: 2000-06-23
PRIORITY APPLICATION NUMBER: US 60/141031
PRIORITY FILING DATE: 1999-06-25
PRIORITY APPLICATION NUMBER: US 60/143262
PRIORITY FILING DATE: 1999-07-09
PRIORITY APPLICATION NUMBER: US 60/151281
PRIORITY FILING DATE: 1999-08-27
PRIORITY APPLICATION NUMBER: DE 19930487.4
PRIORITY FILING DATE: 1999-07-01
PRIORITY APPLICATION NUMBER: DE 19930489.0
PRIORITY FILING DATE: 1999-07-01
PRIORITY APPLICATION NUMBER: DE 19931549.3
PRIORITY FILING DATE: 1999-07-08
PRIORITY APPLICATION NUMBER: DE 19931550.7
PRIORITY FILING DATE: 1999-07-08
PRIORITY APPLICATION NUMBER: DE 19932134.5
PRIORITY FILING DATE: 1999-07-09
PRIORITY APPLICATION NUMBER: DE 19941379.7
PRIORITY FILING DATE: 1999-08-31
Remaining priority application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 446
SEQ ID NO 148
LENGTH: 549
TYPE: PRF
ORGANISM: Corynebacterium glutamicum
US-11-082-389-148

```

[illegible]

Best Local Similarity 25.2%; Pred. No. 7e-23; Matches 149; Conservative 130; Mismatches 234; Indels 79; Gaps 21;

```
QY 977 ISIFLPM--CNHVGALASNYWLSLTDDPIVNGTOEHTKVLASYGALGISQIAVFGYS 1034
Db 247 LSLFLVLVLSLSLEEMALPFFTGRIIDWILQDSADTFRNLTLMSILTTASAVLEF--- 303
QY 1035 MAVSIG-GILASRCLHV-----DLHSILRSPPMSFFERTPSGNLVNRFSEKELPTVDSMI 1087
Db 304 ---VGDIYNNNTGHHVSHLQGEVFGAVLQETEFQONQNGNIMSRVETDSTLSDSL 359
QY 1088 PEVTKMFMGSLFNVIQACIV-IL- ----ATPIAIIIPPLGLYFFVQRFYVASSRQK 1141
Db 360 SENLSTFLMWYL--VRGCLLGIMLMGVSLSLTMVTLITLPLFLPKVKWKYQLLEVOVR 417
QY 1142 RLESVSRSPVYSHNETLIGVSVIRAFEOERFIHQSDLKVDE---NOK---AYYPSIV 1194
Db 418 --ESLAKSSQVA--IEALSMPYTRSPANEGEAKQFREKLQETIKLNQKEAVAY---A 469
QY 1195 ANRW-----IAVRLECVGNCIYLPALPAVISRHSLSAGLVGLSVYSLSQVTTYLNM 1246
Db 470 VNSWTTGISGMLTKVGLIYIGQLVTSGLA-----VSSGNLVTFLYQMQFTQAVEV 520
QY 1247 LVKMSSEMETIVAVERLKEVSETEKEAPMOIQETAPPS---WPQVGRVFPKNYCLRY 1302
Db 521 LLSIYPRVQKAVGSSEKIFEYLDRTPR-----CPPSGLLTPPLHLEGLVQFQDVSPAY 572
QY 1303 --REDLPVLRIHINVTINGEKGIVGRTGAGKSSLTGLFRINSEAGEIIIDGINIAK 1360
Db 573 PNRPDV-LVLOGLFTTLRFGVETALVGPNGSKSTVALNLONLYOPTGGQLLDGKPLPQ 631
QY 1361 IGLHDLRFRKTIIPQDPVLPFSGSLRMLN--DPFSQYSDDEVWTSLELAHLKDFVSALPDK 1418
Db 632 YEHRYLHRQVAAVQEPQVFGRLQENIAYGLTQKPTWEIITAAAVKSGAHSFISGLPQ 691
QY 1419 LDHECARGENLTVGQQLVCLARALLKTKILVIDENATAVD---LETDDLIOSTIRT 1474
Db 692 YDTEVDEAGSQLSGGQRAVALARALIRKPCVLILDDATSAIDANSQLOVEQLYES--P 749
QY 1475 QFEDCTVLTIAHRLNTIMDYTRVYLDKGEIOEVGAPSDILOQRGLFYSMAK 1526
Db 750 ERYRSVLLITQHLSLVEQADHILFLBGGALRBGSTHQQLMEKKGCYAMAVQ 801
```

Search completed: December 15, 2005, 15:50:18
Job time : 12.3405 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:27:22 / Search time 178.197 Seconds
(without alignments)
4800.712 Million cell updates/sec

Title: US-10-665-283-8
Perfect score: 10016
Sequence: 1 MALRGFCASDGLMDMNV.....RSVAVAKKPKFSIPDSLS 1947

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2443163 seqs, 439378781 residues
Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1980s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10016	100.0	1947	9	ADY86945 Human MRP
2	9873	98.6	1927	9	ADY86938 Human MRP
3	9861	98.5	1927	9	ADY86943 MRP1-mous
4	9704	96.9	1891	9	ADY86941 Human MRP
5	7860	78.5	1531	2	AAW57486 Human MRP
6	7860	78.5	1531	2	AAW74471 Human mul
7	7860	78.5	1531	2	AAW98894 Human mul
8	7860	78.5	1531	3	AAW55799 Human mul
9	7860	78.5	1531	3	AAW78873 Human mul
10	7860	78.5	1531	5	ABM35012 Cancer ba
11	7860	78.5	1531	6	ABM35012 Cancer ba
12	7860	78.5	1531	6	ADB20865 MRP1 base
13	7860	78.5	1531	7	ADB87954 Human UGT
14	7860	78.5	1531	7	ADB96937 Human MDR
15	7860	78.5	1531	7	ADB92128 Human MDR
16	7860	78.5	1531	8	ADN44764 Human pro
17	7860	78.5	1531	8	ADN97111 MRP1 prot
18	7860	78.5	1531	8	ADU06426 Novel bro
19	7860	78.5	1531	8	ADU06426 Novel bro
20	7856	78.4	1530	7	ADN95929 Human BEC
21	7849	78.4	1531	2	AAW54928 Multidrug
22	7849	78.4	1531	2	AAW54928 Multidrug
23	7849	78.4	1531	2	AAW57485 Human mul
24	7849	78.4	1531	2	AAW74470 Human mul

25	7849	78.4	1531	2	AAW98893 Human mul
26	7849	78.4	1531	3	AAW55798 Human mul
27	7849	78.4	1531	3	AAW78872 Human mul
28	7849	78.4	1531	3	AAW03582 Multidrug
29	7763	77.5	1515	8	ADU24094 Human cys
30	7610	76.0	1489	2	AAW69952 Multi-dru
31	7587.5	75.8	1482	8	ADN03902 Antipsoi
32	7587.5	75.8	1482	8	ADP23422 PRO poly
33	7572	75.6	1481	5	ABP52108 Homo sapi
34	7560.5	75.2	1482	5	AAW69953 Multi-dru
35	7535	75.2	1475	9	ADW06206 Cyclin-de
36	7510.5	75.0	1472	9	AAW69954 Cyclin-de
37	7500.5	74.9	1472	2	AAW69954 Cyclin-de
38	7230.5	72.2	1530	7	ADP56611 Bovine MR
39	7195.5	71.8	1416	2	ADW06208 Cyclin-de
40	7194	71.8	1417	2	AAW69955 Multi-dru
41	7037	70.3	1388	7	ADJ69912 Human hea
42	7002.5	69.9	1528	2	AAW57487 Murine mu
43	7002.5	69.9	1528	2	AAW74472 Mouse mul
44	7002.5	69.9	1528	2	AAW98895 Mouse mul
45	7002.5	69.9	1528	3	AAW55800 Murine mu

ALIGNMENTS

RESULT 1	ADY86945	ADY86945 standard; protein; 1947 AA.
ID	ADY86945	
AC	ADY86945;	
XX		
DT	02-JUN-2005	(first entry)
XX		
DE	Human MRP1-mouse Klr6.2-HA fusion protein, SEQ ID NO: 8.	
XX		
KW	Ionophore; biosensor; drug screening; diagnostic;	
KW	microorganism detection; potassium channel; fusion protein;	
KW	multidrug resistance protein 1; MRP1; Klr6.2.	
XX		
OS	Homo sapiens.	
OS	Mus musculus.	
OS	Chimeric.	
OS	Unidentified.	
XX		
FT	Key	Location/Qualifiers
FT	Region	1..1531
FT	Region	/note="Multidrug resistance protein 1 (MRP1)"
FT	Region	1532..1537
FT	Region	/note="Hexaglycine spacer"
FT	Region	1538..1947
FT	Region	/note="Klr6.2 protein containing HA epitope peptide"
FT	Region	1651..1659
FT	Region	/note="HA epitope peptide"
XX		
PN	US2005063989-A1.	
XX		
XX	24-MAR-2005.	
XX		
PF	22-SEP-2003; 2003US-00665283.	
XX		
PR	22-SEP-2003; 2003US-00665283.	
XX		
PA	(COMS) COMMISSARIAT ENERGIE ATOMIQUE.	
XX		
PI	Derrand R, Garcia B, Prost A, Revilloud J, Vivaudou M;	
XX	WPI, 2005-252611/26.	
DR		
XX		
PT	New ion channel hybrid protein, used as electrical sensor for screening	
PT	an agonist/antagonist of a membrane protein and for detecting a	
PT	contaminant/pollutant in a sample.	
XX		

PS Claim 25, SEQ ID NO 8, 78bp, English.

XX The present invention relates to a hybrid protein consisting essentially
CC of the fusion of a membrane protein with an ion channel which is not
CC naturally coupled to the membrane protein. The hybrid protein is used as
CC an electrical sensor for screening of an agonist/antagonist of a membrane
CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel Kir6.2 protein containing HA epitope peptide.

XX Sequence 1947 AA:

Query Match 100.0%; Score 10016; DB 9; Length 1947;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1947; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFFLYLSRH 60
DB 1 MALRGFCSADGSDPLMDMNTWNTSNPDFTKCFONTVLVWVPCFYLMACFPFFLYLSRH 60
QY 61 DRGIOMTPLNKKTALGELIMVICMADLFYSFMRSGIFLAPFELVSPFLGITTLLA 120
DB 61 DRGIOMTPLNKKTALGELIMVICMADLFYSFMRSGIFLAPFELVSPFLGITTLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFMLVALCALALIRSKIMTALKEDAQVDLFFDITFYVFS 180
DB 121 TFLIQLERRKGVSSGIMLTFMLVALCALALIRSKIMTALKEDAQVDLFFDITFYVFS 180
QY 181 LLLIQLVLSCPSDRSPLEFSETHIDNPPCRESSASFLSRITTWITGLIVRGYROPLEGGSD 240
DB 181 LLLIQLVLSCPSDRSPLEFSETHIDNPPCRESSASFLSRITTWITGLIVRGYROPLEGGSD 240
QY 241 LMSLNKEDTSQOVPLVVKMKKCECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVVAL 300
DB 241 LMSLNKEDTSQOVPLVVKMKKCECAKTRKQPVKVVYSSKDPAPCKESSKYDANEVVAL 300
QY 301 IVKSPQKEMNPSLFKVLKYTFGPFELMSFFPKAIHDLMFSGPOLIKLLIKFVNDTKAPD 360
DB 301 IVKSPQKEMNPSLFKVLKYTFGPFELMSFFPKAIHDLMFSGPOLIKLLIKFVNDTKAPD 360
QY 361 WQGFYTYVLFLVTAQLQTLVLAHQYFHICFVSGMRITKAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGFYTYVLFLVTAQLQTLVLAHQYFHICFVSGMRITKAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIYNLMSVDAORMDLDATYINMTMSAPLOVITLALYLLMLNGBSVLAVGAVMVLMPVN 480
DB 421 GEIYNLMSVDAORMDLDATYINMTMSAPLOVITLALYLLMLNGBSVLAVGAVMVLMPVN 480
QY 481 AVMAKTKTYOVAAHMSKDNRIKLMNEILNGIKVLKLYAMELAFKDKVLAIROBELKYLK 540
DB 481 AVMAKTKTYOVAAHMSKDNRIKLMNEILNGIKVLKLYAMELAFKDKVLAIROBELKYLK 540
QY 541 KSAVLASAVGTFTWCTPEPLVALCTPAVVYITDENNIIDAOQA FVSLALFNILRPLNITL 600
DB 541 KSAVLASAVGTFTWCTPEPLVALCTPAVVYITDENNIIDAOQA FVSLALFNILRPLNITL 600
QY 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPKXGCGTNSITVRNATTTMARSDPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHEBELPDSIERRPKXGCGTNSITVRNATTTMARSDPT 660
QY 661 LINGITFSLPEGALVAVVGQVCGKSSLLSALLAEWDKVEGHVAKGSVAAYVPOQAMIOND 720
DB 661 LINGITFSLPEGALVAVVGQVCGKSSLLSALLAEWDKVEGHVAKGSVAAYVPOQAMIOND 720
QY 721 SLRENILFEGCOLLEBYRSVIOACALLPDLBILPSGRTIEGKCVNLSSGQKQVSLAR 780
DB 721 SLRENILFEGCOLLEBYRSVIOACALLPDLBILPSGRTIEGKCVNLSSGQKQVSLAR 780
QY 781 AVYGNADLYLEPDDPLSANDAVHGKAI FENVIGPKMKLKKTRILVTHSMSTYLPQVDVITV 840
DB 781 AVYGNADLYLEPDDPLSANDAVHGKAI FENVIGPKMKLKKTRILVTHSMSTYLPQVDVITV 840

QY 841 MSGGKISEMSYQELLARDGAFAEFLRTYASTEQEDQDAEENGVTGVS GPGKAXOMENGM 900
DB 841 MSGGKISEMSYQELLARDGAFAEFLRTYASTEQEDQDAEENGVTGVS GPGKAXOMENGM 900
QY 901 LVYDSAKQIQRQLSSSSSSSGDISRHHNSTAEIQKAEKKEETWKLMEADKQOTGVKL 960
DB 901 LVYDSAKQIQRQLSSSSSSSGDISRHHNSTAEIQKAEKKEETWKLMEADKQOTGVKL 960
QY 961 SVYWDYKKAIGLFFSFSIFLFMCNHSALASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
DB 961 SVYWDYKKAIGLFFSFSIFLFMCNHSALASNYWLSLMTDDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISQGIAVFGYSMAVSIIGIILASRCLAVDLHLSILRSPMSFFERTPSGNLVNRSKEL 1080
DB 1021 ALGISQGIAVFGYSMAVSIIGIILASRCLAVDLHLSILRSPMSFFERTPSGNLVNRSKEL 1080
QY 1081 DTYDSMTPEVITKMPMGLFENVIGACTIYLLATPAAIITPPLGIYFVQRFYVASSROL 1140
DB 1081 DTYDSMTPEVITKMPMGLFENVIGACTIYLLATPAAIITPPLGIYFVQRFYVASSROL 1140
QY 1141 KRLSVSRSPYSHFNETTLGVSIVYRAFEBOERFIHOSDLKVDENQAKYPSIVANRMLA 1200
DB 1141 KRLSVSRSPYSHFNETTLGVSIVYRAFEBOERFIHOSDLKVDENQAKYPSIVANRMLA 1200
QY 1201 VRLCEVNCIYLPALPFAVISRHSLSAGLVGLSVSYSLQYTTYLNMVLVRMSSEMETNIVA 1260
DB 1201 VRLCEVNCIYLPALPFAVISRHSLSAGLVGLSVSYSLQYTTYLNMVLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSERTEKAPMGIOTAPPSWPQYGRVFRNYCLRYRDDLPVLRHINVTINGG 1320
DB 1261 VERLKEYSERTEKAPMGIOTAPPSWPQYGRVFRNYCLRYRDDLPVLRHINVTINGG 1320
QY 1321 EKVGIYVGTGAGKSLTGLFRINESAGEIIDGINIAKIGLHDLKFKITIIPODPVLF 1380
DB 1321 EKVGIYVGTGAGKSLTGLFRINESAGEIIDGINIAKIGLHDLKFKITIIPODPVLF 1380
QY 1381 SGLSRMNLDPFSQYSDSEVWTSLELAHKDFVSALPDKLHBCAEGGEMLSVGOROLVCL 1440
DB 1381 SGLSRMNLDPFSQYSDSEVWTSLELAHKDFVSALPDKLHBCAEGGEMLSVGOROLVCL 1440
QY 1441 ARALIRKTKIIVDEATAVNDLETTDLIOSTIRQOFEDCVYLTIAHLNLTIMYTRIVYL 1500
DB 1441 ARALIRKTKIIVDEATAVNDLETTDLIOSTIRQOFEDCVYLTIAHLNLTIMYTRIVYL 1500
QY 1501 DKGEIOEYGA PSDLLOQRGLFYSMAKXAGLVGGGGGMLSRKGIIPREYVLTIRLAEDPAE 1560
DB 1501 DKGEIOEYGA PSDLLOQRGLFYSMAKXAGLVGGGGGMLSRKGIIPREYVLTIRLAEDPAE 1560
QY 1561 PRYTRRRRARFVSKKNCNVAAHKNIREOGRFLQDVFPTTLVDLTKMPTLLIFTMSFLCSW 1620
DB 1561 PRYTRRRRARFVSKKNCNVAAHKNIREOGRFLQDVFPTTLVDLTKMPTLLIFTMSFLCSW 1620
QY 1621 LLPFAMVWMLIAFAHGDLYAYMEKGITDLAPYDVDPYAGEGTVNVCYTSIHSSSAFLF 1680
DB 1621 LLPFAMVWMLIAFAHGDLYAYMEKGITDLAPYDVDPYAGEGTVNVCYTSIHSSSAFLF 1680
QY 1681 SIEVOVITIGRCGRWTEBCPLAILILIVONIVGMINAIVMGCIFMTAQAHRAETLIF 1740
DB 1681 SIEVOVITIGRCGRWTEBCPLAILILIVONIVGMINAIVMGCIFMTAQAHRAETLIF 1740
QY 1741 SKHAVITLRHGRLCFMLRVGDLRKSMISATIHQVVRKTTSPSGEVVPLAHQVIDIPENG 1800
DB 1741 SKHAVITLRHGRLCFMLRVGDLRKSMISATIHQVVRKTTSPSGEVVPLAHQVIDIPENG 1800
QY 1801 VGGNGIFLVAPLIIYHYIDSNSPLYDLAPSLDHHODLEIIVIEGVEVETGITTQARTS 1860
DB 1801 VGGNGIFLVAPLIIYHYIDSNSPLYDLAPSLDHHODLEIIVIEGVEVETGITTQARTS 1860
QY 1861 YLADEILMGRFVYIVAEEDRYSVDYSKRGNTIKVPTPLCTARQLEDSDSLALTLAS 1920
DB 1861 YLADEILMGRFVYIVAEEDRYSVDYSKRGNTIKVPTPLCTARQLEDSDSLALTLAS 1920

QY 1921 SRGFLRKRSVAVAKAKPKFSISPDLS 1947
 DB 1921 SRGFLRKRSVAVAKAKPKFSISPDLS 1947

RESULT 2
 ID ADY86938 standard; protein; 1927 AA.
 AC ADY86938;
 XX
 DT 02-JUN-2005 (first entry)
 DE Human MRP1-mouse Klr6.2 fusion protein, SEQ ID NO: 1.
 XX
 KM Ionophore; biosensor; drug screening; diagnostic;
 KM microorganism detection; potassium channel; fusion protein;
 KM multidrug resistance protein 1; MRP1; Klr6.2.
 XX Homo sapiens.
 OS Mus musculus.
 OS Chimeric.
 XX
 XX Key Location/Qualifiers
 FH 1. .1531
 FT /note= "Multidrug resistance protein 1 (MRP1) "
 FT Region 1532. .1537
 FT /note= "Hexaglycine spacer"
 FT 1538. .1927
 FT Region /note= "Klr6.2 protein"
 PN US2005063989-A1.
 XX
 PD 24-MAR-2005.
 XX
 PF 22-SEP-2003; 2003US-00665283.
 XX
 PR 22-SEP-2003; 2003US-00665283.
 XX
 PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
 PI Derand R, Garcia E, Prost A, Revilloud J, Vivaudou M;
 DR WPI; 2005-252611/26.
 XX
 PT New ion channel hybrid protein, used as electrical sensor for screening
 PT an agonist/antagonist of a membrane protein and for detecting a
 PT contaminant/pollutant in a sample.
 XX
 PS Claim 25; SEQ ID NO 1; 78pp; English.
 XX
 CC The present invention relates to a hybrid protein consisting essentially
 CC of the fusion of a membrane protein with an ion channel which is not
 CC naturally coupled to the membrane protein. The hybrid protein is used as
 CC an electrical sensor for screening of an agonist/antagonist of a membrane
 CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
 CC invention is also useful for medical diagnostics and microorganism
 CC detection. The present sequence is a fusion protein comprising human
 CC multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
 CC channel Klr6.2 protein.
 CC
 XX
 SQ Sequence 1927 AA;
 Query Match 98.6%; Score 9873; DB 9; Length 1927;
 Best Local Similarity 99.0%; Pired. No. 0;
 Matches 1927; Conservative 0; Mismatches 0; Indels 20; Gaps 2;
 QY 1 MALRGFCSADGSPILMDMNTANTSNPDFKCFONTLVVWPCFYLMACPFYFLYSRH 60
 DB 1 MALRGFCSADGSPILMDMNTANTSNPDFKCFONTLVVWPCFYLMACPFYFLYSRH 60
 QY DRGYIQMTPLNKTKTALGFLIMIVCWADLFYFWMERSRGIFLAPVFLVSPTLIGITTLA 120

DB 61 DRGYIQMTPLNKTKTALGFLIMIVCWADLFYFWMERSRGIFLAPVFLVSPTLIGITTLA 120
 QY 121 TPLIQLERRRGVSSGIMLTFWVLVYCALAIIIRSKITMTALKEAQVDLFRDITFYFYFS 180
 DB 121 TPLIQLERRRGVSSGIMLTFWVLVYCALAIIIRSKITMTALKEAQVDLFRDITFYFYFS 180
 QY 181 LLLIQVLVSCFSPRSLPFSETHDPNCPSSSASFISRIITFWMTGIIIVGYRPLGSD 240
 DB 181 LLLIQVLVSCFSPRSLPFSETHDPNCPSSSASFISRIITFWMTGIIIVGYRPLGSD 240
 QY 241 LMSLNKEDTSEQVVPVLVKNWKKECAKTRKOPVKVYSSKDPAPCKESSKVDANEVEAL 300
 DB 241 LMSLNKEDTSEQVVPVLVKNWKKECAKTRKOPVKVYSSKDPAPCKESSKVDANEVEAL 300
 QY 301 IVKSPQKEMNPSPFKVLKTFPGPYFLMSFFPKAIHDLMSGPOILKLIKFNVDTPAD 360
 DB 301 IVKSPQKEMNPSPFKVLKTFPGPYFLMSFFPKAIHDLMSGPOILKLIKFNVDTPAD 360
 QY 361 MOGYFYTVLLFVTRACIQTLLVHOYFHCIVSGMRKIKTAIVGAYRKALVITNSARKSTV 420
 DB 361 MOGYFYTVLLFVTRACIQTLLVHOYFHCIVSGMRKIKTAIVGAYRKALVITNSARKSTV 420
 QY 421 GEIVNIMSVDQRFMDLATYINMIWSAPLOVILALYLMNLGSPVLAGVAVMYLAMPVN 480
 DB 421 GEIVNIMSVDQRFMDLATYINMIWSAPLOVILALYLMNLGSPVLAGVAVMYLAMPVN 480
 QY 481 AVNARKTKTYQVAHMSKDNRIKLANEILNGIKVLKIYAMELAPKOKVLAIROBELKVLK 540
 DB 481 AVNARKTKTYQVAHMSKDNRIKLANEILNGIKVLKIYAMELAPKOKVLAIROBELKVLK 540
 QY 541 KSAVYLSAVGFTFWCTPPELVALCTPAVYVITIDNNIILDQTAIVSIALFNIILPPLNLP 600
 DB 541 KSAVYLSAVGFTFWCTPPELVALCTPAVYVITIDNNIILDQTAIVSIALFNIILPPLNLP 600
 QY 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKGSGTNSITYNATFTMARSPPT 660
 DB 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKGSGTNSITYNATFTMARSPPT 660
 QY 661 LMGITTSIPRGALVAVVGQVCGKSSILSLALAEKMKVEGHVAIKGSVAVYPOQAWIOND 720
 DB 661 LMGITTSIPRGALVAVVGQVCGKSSILSLALAEKMKVEGHVAIKGSVAVYPOQAWIOND 720
 QY 721 SLRENIILFGQLEPEPYRSVIOACALLPDEILIPSGDREITGRKYNLSSGQORVSLAR 780
 DB 721 SLRENIILFGQLEPEPYRSVIOACALLPDEILIPSGDREITGRKYNLSSGQORVSLAR 780
 QY 781 AVYSNADIVLFDPLPLAVDAVDAVGHKILFENVYIGPKMKLNKTRILVTHSMGYLPOVDYIV 840
 DB 781 AVYSNADIVLFDPLPLAVDAVDAVGHKILFENVYIGPKMKLNKTRILVTHSMGYLPOVDYIV 840
 QY 841 MSGGKISEMGSYQELLARDGAFAPLRTYASTBOEDAEENGVTGVSGPKKAKOMENGM 900
 DB 841 MSGGKISEMGSYQELLARDGAFAPLRTYASTBOEDAEENGVTGVSGPKKAKOMENGM 900
 QY 901 LVTDSAGKQORLSSSSSVSGDISRHNSHTALQKAEAKKETWLMEMDAKQOTQVYL 960
 DB 901 LVTDSAGKQORLSSSSSVSGDISRHNSHTALQKAEAKKETWLMEMDAKQOTQVYL 960
 QY 961 SYVWDYMKAIQLFISPLSIFLFWCNHVSALASAYMLSLMTDDIYVNGTOBHTKRLSVYG 1020
 DB 961 SYVWDYMKAIQLFISPLSIFLFWCNHVSALASAYMLSLMTDDIYVNGTOBHTKRLSVYG 1020
 QY 1021 ALGISQGIAVFGYSMAVSIIGIILASRCLHVDLHSLIRSPMSFERTPSGNLVNRFSEL 1080
 DB 1021 ALGISQGIAVFGYSMAVSIIGIILASRCLHVDLHSLIRSPMSFERTPSGNLVNRFSEL 1080
 QY 1081 DTVDSMIPVYIKMFMKSLFNVIGACTVILLATPILAIIPGLIYFVFORFVYASSROU 1140
 DB 1081 DTVDSMIPVYIKMFMKSLFNVIGACTVILLATPILAIIPGLIYFVFORFVYASSROU 1140
 QY 1141 KRLSEVSRSPVYSHFNETLLGVSVIRAFEOERFIHQSLDKVDENOKAYYPSIVANRWLA 1200
 DB 1141 KRLSEVSRSPVYSHFNETLLGVSVIRAFEOERFIHQSLDKVDENOKAYYPSIVANRWLA 1200

```

QY 1201 VRLECVGNCLVLFALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSSEMETNIVA 1260
Db 1201 VRLECVGNCLVLFALFAVISRHSLSAGLVGLSVSYSLQVTTYLNMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSFETKEAPMOIQETAPPSWPQVRVEFRNYCLARYEDLPVLAHINVTINGG 1320
Db 1261 VERLKEYSFETKEAPMOIQETAPPSWPQVRVEFRNYCLARYEDLPVLAHINVTINGG 1320
QY 1321 EKVGIVERTGAGKSLTGLFRINESARGEIITIGINAKIGLHDLRKITITIIQODPLF 1380
Db 1321 EKVGIVERTGAGKSLTGLFRINESARGEIITIGINAKIGLHDLRKITITIIQODPLF 1380
QY 1381 SGLSRMNLDPFSSQYSDSEWMTSLBLAHLKDFVSALPDKLDHECAEGENLSVGQRLVCL 1440
Db 1381 SGLSRMNLDPFSSQYSDSEWMTSLBLAHLKDFVSALPDKLDHECAEGENLSVGQRLVCL 1440
QY 1441 ARALLRTKILVDEBATAVDLETDLLQSTIRQFEDCTVLTIAHRLNTIMDTRYIVL 1500
Db 1441 ARALLRTKILVDEBATAVDLETDLLQSTIRQFEDCTVLTIAHRLNTIMDTRYIVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYMAKADAGLVGGGGGMLSRKGIIPBEYVLTPLAEDPAE 1560
Db 1501 DKGEIOEYGAPSDLLQORGLFYMAKADAGLVGGGGGMLSRKGIIPBEYVLTPLAEDPAE 1560
QY 1561 PRYTRRRARFVSKGNCNVAHKNIROGRFLQDVFTTLYDLKMPHTLLIFTMSFLCSW 1620
Db 1561 PRYTRRRARFVSKGNCNVAHKNIROGRFLQDVFTTLYDLKMPHTLLIFTMSFLCSW 1620
QY 1621 LLEFAMWWMLIAFAHGDLVAYMEKGTDLAPYRVDYDVAAGSTNVPCYTSIHSFSALEF 1680
Db 1621 LLEFAMWWMLIAFAHGDLVAYMEKGTDLAPYRVDYDVAAGSTNVPCYTSIHSFSALEF 1680
QY 1681 SIEVOVTIGFGGMYTECPALILILYQNTVGLMINAIMGCIPIKTAQAARRAETLIIF 1740
Db 1681 SIEVOVTIGFGGMYTECPALILILYQNTVGLMINAIMGCIPIKTAQAARRAETLIIF 1740
QY 1741 SKHAVITLHRGLCFMRLRVGDLRKSMIISATIHQVVKTTSPBEGVVPPLHQVDPIMENG 1800
Db 1741 SKHAVITLHRGLCFMRLRVGDLRKSMIISATIHQVVKTTSPBEGVVPPLHQVDPIMENG 1800
QY 1801 VGGNGIFLVAPLIIVYHVDNSPLVDLAPSDLHHQDLEIIVLEGVETGITTQATS 1860
Db 1801 VGGNGIFLVAPLIIVYHVDNSPLVDLAPSDLHHQDLEIIVLEGVETGITTQATS 1860
QY 1861 YLADEIILMGORFPIVABEDGRYSVDYSGFNTIKVPTPLCTAROLDERSILDTLAS 1920
Db 1861 YLADEIILMGORFPIVABEDGRYSVDYSGFNTIKVPTPLCTAROLDERSILDTLAS 1920
QY 1921 SRGPIRKRSVAVAKAKPKFSISPDSLS 1947
Db 1921 SRGPIRKRSVAVAKAKPKFSISPDSLS 1947
QY 1901 SRGPIRKRSVAVAKAKPKFSISPDSLS 1927
Db 1901 SRGPIRKRSVAVAKAKPKFSISPDSLS 1927

RESULT 3
ADY86943
ID ADY86943 standard; protein; 1927 AA.
XX
AC ADY86943;
DT 02-JUN-2005 (first entry)
DE MRP1-mouse Klr6.2 mutant (KR370AA) fusion protein, SEQ ID NO: 6.
DM
KM Ionophore; biosensor; drug screening; diagnostic;
KW microorganism detection; potassium channel; fusion protein;
KM multidrug resistance protein 1; MRP1; Klr6.2; muttein.
XX
OS Homo sapiens.
OS Mus musculus.
OS Chimeric.
XX
FH Key Location/Qualifiers

```

```

FT Region
FT 1. .1531
FT /note= "Multidrug resistance protein 1 (MRP1)"
FT Region
FT 1532. .1537
FT /note= "Hexaglycine spacer"
FT Region
FT 1538. .1927
FT /note= "Klr6.2 mutant protein"
FT Misc-difference
FT 1907
FT /note= "wild-type lys substituted by Ala"
FT Misc-difference
FT 1908
FT /note= "wild-type Arg substituted by Ala"
FT US2005063989-A1.
FT 24-MAR-2005.
FT 22-SEP-2003; 2003US-00665283.
FT 22-SEP-2003; 2003US-00665283.
FT 22-SEP-2003; 2003US-00665283.
FT (COMS ) COMMISSARIAT ENERGIE ATOMIQUE.
FT Derand R, Garcia E, Prost A, Revilloud J, Vivaudou M;
FT WPI; 2005-252611/26.
FT
FT New ion channel hybrid protein, used as electrical sensor for screening
FT an agonist/antagonist of a membrane protein and for detecting a
FT contaminant/pollutant in a sample.
FT
FT Claim 25; SEQ ID NO 6; 78bp; English.
XX
XX The present invention relates to a hybrid protein consisting essentially
XX of the fusion of a membrane protein with an ion channel which is not
XX naturally coupled to the membrane protein. The hybrid protein is used
XX as an electrical sensor for screening of an agonist/antagonist of a membrane
XX protein, drugs and for detecting a contaminant/pollutant in a sample. The
XX invention is also useful for medical diagnostics and microorganism
XX detection. The present sequence is a fusion protein comprising human
XX multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
XX channel Klr6.2 mutant (KR370AA) protein.
XX
XX Sequence 1927 AA.
SQ
Query Match 98.5%; Score 9861; DB 9; Length 1927;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1925; Conservative 0; Mismatches 2; Indels 20; Gaps 2;
QY 1 MALRGFSGADGSDPLMDMNTWNTNSNDFTKCFQNTYLVWVPCRYLWACFPFFLYLSRH 60
Db 1 MALRGFSGADGSDPLMDMNTWNTNSNDFTKCFQNTYLVWVPCRYLWACFPFFLYLSRH 60
QY 61 DRGYIOMTPLUNKKTALGFLIMIVCMADLFYSPFERSNGIPLAVPLVSPTLGITTLLA 120
Db 61 DRGYIOMTPLUNKKTALGFLIMIVCMADLFYSPFERSNGIPLAVPLVSPTLGITTLLA 120
QY 121 TFLIQLERRKGVQSSGIMLTFWLVALVCALALIRSKIMTALKEDAQVDLFPDITFYVYS 180
Db 121 TFLIQLERRKGVQSSGIMLTFWLVALVCALALIRSKIMTALKEDAQVDLFPDITFYVYS 180
QY 181 LLLIQLVISCFSRSPFLFSETHDNPNCBSSASFLSRITRWITGLIVRGYRQPLBSSD 240
Db 181 LLLIQLVISCFSRSPFLFSETHDNPNCBSSASFLSRITRWITGLIVRGYRQPLBSSD 240
QY 241 LMSLNKEDTSEQVVPVIVKMKKCAKTRQPVVYVSSQDPAQPKSSKYDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVIVKMKKCAKTRQPVVYVSSQDPAQPKSSKYDANEVEAL 300
QY 301 IVKSPQKEMNPSLRFVLYKTFGPFYLMSPFKAIHDLMMFSGPOLIKLLIKFVNDTKAPD 360
Db 301 IVKSPQKEMNPSLRFVLYKTFGPFYLMSPFKAIHDLMMFSGPOLIKLLIKFVNDTKAPD 360
QY 361 WQGYFFYVLLFVTAQLOTLVIHQYFHIQFVSGMRITKAVIGAVYRKALVTNSAKSSTV 420

```

Db 361 WOGFYVTLVFTVACIOTLVHQYFHCIFVSGMRKTAIVAGVYRKALVITNSARKSSTV 420
 Qy 421 GETVNLMSVDAORPMDLATYINMWSAFLQVYTLALYLLMLNGSVTLAGVAVMLAMPVN 480
 Db 421 GETVNLMSVDAORPMDLATYINMWSAFLQVYTLALYLLMLNGSVTLAGVAVMLAMPVN 480
 Qy 481 AVAMKTKTYOVAMHMSKDNRIKLMELINGIKVLKLYAMELAFKDYKLYI ROBELVYK 540
 Db 481 AVAMKTKTYOVAMHMSKDNRIKLMELINGIKVLKLYAMELAFKDYKLYI ROBELVYK 540
 Qy 541 KSAVLSAVGFTFWCTPFLVALCTFAVYVITDENNNILDAQTAFAVSLFNLRLPPLNLP 600
 Db 541 KSAVLSAVGFTFWCTPFLVALCTFAVYVITDENNNILDAQTAFAVSLFNLRLPPLNLP 600
 Qy 601 MVSIVQASVSLKRLIFLSHELEPDSIERRPVKGCGNSTYVNAATFTMARSPPT 660
 Db 601 MVSIVQASVSLKRLIFLSHELEPDSIERRPVKGCGNSTYVNAATFTMARSPPT 660
 Qy 661 LMGITFSPGALVAVGVGCGKSSLSALLAEMDVEGHVAIKGSVAIVPOQAMIOND 720
 Db 661 LMGITFSPGALVAVGVGCGKSSLSALLAEMDVEGHVAIKGSVAIVPOQAMIOND 720
 Qy 721 SLRENILFGCQLEBPYRSVIOACALLPDEILPSGDRTEIGEKVNLSSGQKQVSLAR 780
 Db 721 SLRENILFGCQLEBPYRSVIOACALLPDEILPSGDRTEIGEKVNLSSGQKQVSLAR 780
 Qy 781 AYSNNDIYLPDPLSAVDHVGKHIFENYIGPKGMKNTKRIILVTHSMYTLQVDYIIV 840
 Db 781 AYSNNDIYLPDPLSAVDHVGKHIFENYIGPKGMKNTKRIILVTHSMYTLQVDYIIV 840
 Qy 841 MSGGKISEMGSYOELARDAFAPLRTYASTBOEBAEENGVTGVSGPKKAKOMENGM 900
 Db 841 MSGGKISEMGSYOELARDAFAPLRTYASTBOEBAEENGVTGVSGPKKAKOMENGM 900
 Qy 901 LVYDSAGKQLOROLSSSSSYSGDISRHNSSTAELOKAEKKEBTWKLEADKAQTQGVKL 960
 Db 901 LVYDSAGKQLOROLSSSSSYSGDISRHNSSTAELOKAEKKEBTWKLEADKAQTQGVKL 960
 Qy 961 SVYWDYMKALGLTSLSTFLFMCNHYSAIASNYWLSLMTDDPIVNGTOHTYKRLSVYG 1020
 Db 961 SVYWDYMKALGLTSLSTFLFMCNHYSAIASNYWLSLMTDDPIVNGTOHTYKRLSVYG 1020
 Qy 1021 ALGISOGIAVFGYSMAVISGIIILASRCLAHVDLHSIRSPMSFEERTPSGNLNVRSFKEL 1080
 Db 1021 ALGISOGIAVFGYSMAVISGIIILASRCLAHVDLHSIRSPMSFEERTPSGNLNVRSFKEL 1080
 Qy 1081 DTVDMSIPEVIKMFMSGLFNVIACIYILATPIAIIIPPLGIYFFVORFYVASSROL 1140
 Db 1081 DTVDMSIPEVIKMFMSGLFNVIACIYILATPIAIIIPPLGIYFFVORFYVASSROL 1140
 Qy 1141 KRLSVSRSPYGHFNMTLLGVSIVIRAFEOERFIHOSDLKVDENOKAYPISIVANKWLA 1200
 Db 1141 KRLSVSRSPYGHFNMTLLGVSIVIRAFEOERFIHOSDLKVDENOKAYPISIVANKWLA 1200
 Qy 1201 VRLCEGNCIVLPAALPAVISRHSLSAGVLSVSYSLQVTTYLTANMLVYRMSSEMETNIVA 1260
 Db 1201 VRLCEGNCIVLPAALPAVISRHSLSAGVLSVSYSLQVTTYLTANMLVYRMSSEMETNIVA 1260
 Qy 1261 VERLKEYSETEKAPMOIQETAPSSWPQGVREFRNYCLYREDDLDVLRHINVTINGG 1320
 Db 1261 VERLKEYSETEKAPMOIQETAPSSWPQGVREFRNYCLYREDDLDVLRHINVTINGG 1320
 Qy 1321 EKVGIVGRGAGKSSLTGLFRINESABEIIIDGINIAKIGLHDLRFKTTIIPDPVLF 1380
 Db 1321 EKVGIVGRGAGKSSLTGLFRINESABEIIIDGINIAKIGLHDLRFKTTIIPDPVLF 1380
 Qy 1381 SCSLRNMLDPFSQYDEEVTSLLELAHLKDFVSLPDKLHCEAEGENSVGORGVLCL 1440
 Db 1381 SCSLRNMLDPFSQYDEEVTSLLELAHLKDFVSLPDKLHCEAEGENSVGORGVLCL 1440
 Qy 1441 ARALLKTKTILVDEATAVDLETDLIQSTIRTOFEDCTVLTARHLANTIMDYTRVYL 1500
 Db 1441 ARALLKTKTILVDEATAVDLETDLIQSTIRTOFEDCTVLTARHLANTIMDYTRVYL 1500

Qy 1501 DKGEIOEYGA PSDLLQORGLFYSMAKDAGLVGGGGGGMLSRKGIIPREYVLTPLAEPDAE 1560
 Db 1501 DKGEIOEYGA PSDLLQORGLFYSMAKDAGLVGGGGGGMLSRKGIIPREYVLTPLAEPDAE 1560
 Qy 1561 PRYTERRRARFVSKKGNCAVAKHNIREQRFLODYFTTLVDLKWPHLLIPTMSFLCSW 1620
 Db 1561 PRYTERRRARFVSKKGNCAVAKHNIREQRFLODYFTTLVDLKWPHLLIPTMSFLCSW 1620
 Qy 1621 LFLPMVWMLIAPAHGDIYAYMEKGITDLADYPYDVPDYAGEGTVNVCVTSIHSSSAFLF 1680
 Db 1621 LFLPMVWMLIAPAHGDIYAYMEKGITDLADYPYDVPDYAGEGTVNVCVTSIHSSSAFLF 1680
 Qy 1681 SIEVOVTIGRGGMWTEECPLATILLIVQNVGLMNTAMLMGCI PMKTAQAHRAETLIF 1740
 Db 1681 SIEVOVTIGRGGMWTEECPLATILLIVQNVGLMNTAMLMGCI PMKTAQAHRAETLIF 1740
 Qy 1741 SKHAVITLRRGRCLCFMLRVGDLRKSMIISATIMQVVRKTSPEGEVVPPLHQVDIPMENG 1800
 Db 1741 SKHAVITLRRGRCLCFMLRVGDLRKSMIISATIMQVVRKTSPEGEVVPPLHQVDIPMENG 1800
 Qy 1781 VGGNGIFLVAPLLIYHVIDSNSPLYDLAPSDLHHODLEIIVILGVEVETGITTQARTS 1860
 Db 1781 VGGNGIFLVAPLLIYHVIDSNSPLYDLAPSDLHHODLEIIVILGVEVETGITTQARTS 1860
 Qy 1861 YLADELIMGQRFPIVAEBDGRYSVYSKRGNTIKVPTPLCTAROLDSDSLDALTLAS 1920
 Db 1861 YLADELIMGQRFPIVAEBDGRYSVYSKRGNTIKVPTPLCTAROLDSDSLDALTLAS 1920
 Qy 1921 SRGPLRKRSAVAKAKPKFSISPDLS 1947
 Db 1901 SRGPLRAASA VAVAKAKPKFSISPDLS 1927

RESULT 4
 ADY86941
 ID ADY86941 standard; protein; 1891 AA.
 AC ADY86941;
 AC
 DT 02-JUN-2005 (first entry)
 XX
 DE Human MRP1-kir6.2 mutant delctac36 fusion protein, SRQ ID NO: 4.
 XX
 KM Ionophore; biosensor; drug screening; diagnostic;
 KM microorganism detection; potassium channel; fusion protein;
 KM multidrug resistance protein 1; MRP1; Kir6.2; mutuin.
 OS Homo sapiens.
 OS Mus musculus.
 OS Chimeric.
 OS
 XX
 FH Key
 FT Region 1..1531
 FT /note= "Multidrug resistance protein 1 (MRP1)"
 FT Region 1532..1537
 FT /note= "Hexaglycine spacer"
 FT Region 1538..1891
 FT /note= "Kir6.2 mutant delctac36 protein"
 PN US2005063989-A1.
 XX
 PD 24-MAR-2005.
 XX
 PF 22-SEP-2003; 2003US-00665283.
 XX
 PR 22-SEP-2003; 2003US-00665283.
 XX
 PA (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
 XX
 PI Derand R, Garcia E, Prost A, Revilloud J, Viaudou M;
 XX
 DR WPI; 2005-252611/26.

XX New ion channel hybrid protein, used as electrical sensor for screening
PT an agonist/antagonist of a membrane protein and for detecting a
PT contaminant/pollutant in a sample.
XX
PS Claim 25; SEQ ID NO 4; 78pp; English.
XX
CC The present invention relates to a hybrid protein consisting essentially
CC of the fusion of a membrane protein with an ion channel which is not
CC naturally coupled to the membrane protein. The hybrid protein is used as
CC an electrical sensor for screening of an agonist/antagonist of a membrane
CC protein, drugs and for detecting a contaminant/pollutant in a sample. The
CC invention is also useful for medical diagnostics and microorganism
CC detection. The present sequence is a fusion protein comprising human
CC multidrug resistance protein 1 (MRP1) and mouse inner rectifier potassium
CC channel Kir6.2 mutant detac36 protein.
XX
SQ Sequence 1891 AA;
Query Match 96.9%; Score 9704; DB 9; Length 1891;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1891; Conservative 0; Mismatches 0; Indels 20; Gaps 2;
QY 1 MALRGFASDGDPLMDMNTWNTSNPDFTKCFQNTLVWVPCGYLWACPFYLYSRH 60
DB 1 MALRGFASDGDPLMDMNTWNTSNPDFTKCFQNTLVWVPCGYLWACPFYLYSRH 60
QY DRGYIQMTPLNKTALGFLIMIVCMADLFYSPWERSGIFLAVFLVSPILGITTLLA 120
DB DRGYIQMTPLNKTALGFLIMIVCMADLFYSPWERSGIFLAVFLVSPILGITTLLA 120
QY 61 DRGYIQMTPLNKTALGFLIMIVCMADLFYSPWERSGIFLAVFLVSPILGITTLLA 120
DB 61 DRGYIQMTPLNKTALGFLIMIVCMADLFYSPWERSGIFLAVFLVSPILGITTLLA 120
QY 121 TPLQLERRKGVSGSGLTFMWLVVALCALILRSKIMTALKEDAQYDLFEDITFYVYS 180
DB 121 TPLQLERRKGVSGSGLTFMWLVVALCALILRSKIMTALKEDAQYDLFEDITFYVYS 180
QY 181 LLILQLVLSGSDSPSLFSETHDNPBPRESSAFLSRITFWMTGLIVRGYRPLGSD 240
DB 181 LLILQLVLSGSDSPSLFSETHDNPBPRESSAFLSRITFWMTGLIVRGYRPLGSD 240
QY 241 LMSLNKEDTSEQVAVLVKMKKECAKTRKQPVKVVSSKDPAPKSSKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVAVLVKMKKECAKTRKQPVKVVSSKDPAPKSSKVDANEVEAL 300
QY 301 IVSPQKEMNPSLFKVLKYTFGPFLMSFFPKAIDLMFSGPOLKLLIKFVNDTRAP 360
DB 301 IVSPQKEMNPSLFKVLKYTFGPFLMSFFPKAIDLMFSGPOLKLLIKFVNDTRAP 360
QY 361 WQGYFTYVLLFTVACLOTLVHOYFHICFVSGMKIKTAVIGAVYRKALVITNSARKSTV 420
DB 361 WQGYFTYVLLFTVACLOTLVHOYFHICFVSGMKIKTAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIYNLMSVAORFMDLATYINMISAPLOVTLALYLLMLNIGPSVLAVGVAWVLAMPVN 480
DB 421 GEIYNLMSVAORFMDLATYINMISAPLOVTLALYLLMLNIGPSVLAVGVAWVLAMPVN 480
QY 481 AVAMAKTKTYOVAMKSKDNRIKLMNEILNGIKVLYKUYAMELAFKDKVLAIRQELKVLK 540
DB 481 AVAMAKTKTYOVAMKSKDNRIKLMNEILNGIKVLYKUYAMELAFKDKVLAIRQELKVLK 540
QY 541 KSAVLAVGFTTWCTPPLVALCTPFAVYVITDENNIIDAQTAFLSLFNLIRPPLNLP 600
DB 541 KSAVLAVGFTTWCTPPLVALCTPFAVYVITDENNIIDAQTAFLSLFNLIRPPLNLP 600
QY 601 MVISSIVQASVSLKRLRILPSHEELPDSITERRPVKGGGNSITVNAATTYARSDPT 660
DB 601 MVISSIVQASVSLKRLRILPSHEELPDSITERRPVKGGGNSITVNAATTYARSDPT 660
QY 661 LMGITFSPREGALVAVVGVCGCKSLSLALIAEMDKVEGHVAKGSVAVYVPOQAMIQND 720
DB 661 LMGITFSPREGALVAVVGVCGCKSLSLALIAEMDKVEGHVAKGSVAVYVPOQAMIQND 720
QY 721 SLRENILFGCOLLEPPYRSVTOACALLPDLIELPSGDRTEIGEKVNLSSGQKORVSLAR 780
DB 721 SLRENILFGCOLLEPPYRSVTOACALLPDLIELPSGDRTEIGEKVNLSSGQKORVSLAR 780

DB 721 SLRENILFGCOLLEPPYRSVTOACALLPDLIELPSGDRTEIGEKVNLSSGQKORVSLAR 780
QY 781 AVYSNADITYLFDPLSLAVDAVGHKILFENVIGRGMKNKTRILVTHSMGYLPQVDIYV 840
DB 781 AVYSNADITYLFDPLSLAVDAVGHKILFENVIGRGMKNKTRILVTHSMGYLPQVDIYV 840
QY 841 MSGGKISEMGSYOBLLARDGAFAPLRTYASTQEOBDAENGVTVGSGPGEKAKOMENGM 900
DB 841 MSGGKISEMGSYOBLLARDGAFAPLRTYASTQEOBDAENGVTVGSGPGEKAKOMENGM 900
QY 901 LVNDSAGKOLQORLSSSSSYSGDISRHNSHTALQKAEKEETWKLMEADKQOTGVKL 960
DB 901 LVNDSAGKOLQORLSSSSSYSGDISRHNSHTALQKAEKEETWKLMEADKQOTGVKL 960
QY 961 SYVDWYKAIQGLIFSLFSLFPMCNHVSALASNYMLSLMTDDPIVNGTOHTKVRSLSVYG 1020
DB 961 SYVDWYKAIQGLIFSLFSLFPMCNHVSALASNYMLSLMTDDPIVNGTOHTKVRSLSVYG 1020
QY 1021 ALGISGIAVFGYSMAVSIIGIILASRCLHYDLHSILRSPMSFPERTSGNLVNRSEKEL 1080
DB 1021 ALGISGIAVFGYSMAVSIIGIILASRCLHYDLHSILRSPMSFPERTSGNLVNRSEKEL 1080
QY 1081 DTVDGMTPEYIKPMFMSLFENVIGACIVTILATPIAIIIPPLGLIYFVQRFVYASSROL 1140
DB 1081 DTVDGMTPEYIKPMFMSLFENVIGACIVTILATPIAIIIPPLGLIYFVQRFVYASSROL 1140
QY 1141 KRLESVSRSPVYSHFNETLIGSVIRAFEBQERFIHOSDLKVDENQKAYVPSIVANRWLA 1200
DB 1141 KRLESVSRSPVYSHFNETLIGSVIRAFEBQERFIHOSDLKVDENQKAYVPSIVANRWLA 1200
QY 1201 VRLBVGNCIVLPAALFAVISRHSLSAGLVGSVSLQVTTYLNLVRRMSSEMETNIVA 1260
DB 1201 VRLBVGNCIVLPAALFAVISRHSLSAGLVGSVSLQVTTYLNLVRRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMQIOETAPSSMPQVGRVFRFRYCLARYEDDLFVLRHINVTNGG 1320
DB 1261 VERLKEYSETEKAPMQIOETAPSSMPQVGRVFRFRYCLARYEDDLFVLRHINVTNGG 1320
QY 1321 EKVYIGVRTAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLV 1380
DB 1321 EKVYIGVRTAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLV 1380
QY 1381 SGIIRMLNLDPPSQYSDERWTSLELAHLKDFVSAIPKLDHBCAEGGENTSVGQROLVCL 1440
DB 1381 SGIIRMLNLDPPSQYSDERWTSLELAHLKDFVSAIPKLDHBCAEGGENTSVGQROLVCL 1440
QY 1441 ARALARTKTLVDEATAVADLETDDLIOSTIRQFEDCTVLITAHRLNTIMOTRYIVL 1500
DB 1441 ARALARTKTLVDEATAVADLETDDLIOSTIRQFEDCTVLITAHRLNTIMOTRYIVL 1500
QY 1501 DKGBIOEYGAAPSDILOQRGLFYSMADKAGLVGGGGGMLSRKGIIPREYVLTRLAEPPAE 1560
DB 1501 DKGBIOEYGAAPSDILOQRGLFYSMADKAGLVGGGGGMLSRKGIIPREYVLTRLAEPPAE 1560
QY 1561 PRYTRERRARFVSKKNCNVAAKNIREQGRPIQDVFTLLVLDKWPHTLLIFTMSFLCSW 1620
DB 1561 PRYTRERRARFVSKKNCNVAAKNIREQGRPIQDVFTLLVLDKWPHTLLIFTMSFLCSW 1620
QY 1621 LLPAMVWMLLAPAHGDIYAMKEGIDPLAIPYDPVPAABEGTNPVCVTSIHSSSAFLF 1680
DB 1621 LLPAMVWMLLAPAHGDIYAMKEGIDPLAIPYDPVPAABEGTNPVCVTSIHSSSAFLF 1680
QY 1681 SIEVQVITIGFGNNVTECEPLAILILVONIVGLMINAIMGCIEMTQAQHRRAETLIF 1740
DB 1681 SIEVQVITIGFGNNVTECEPLAILILVONIVGLMINAIMGCIEMTQAQHRRAETLIF 1740
QY 1741 SKNAVITLRHGRCLCFMLRVGDLKRSMLISATIHQVVRKTTSPBGEVVLPHQVDIPENG 1800
DB 1741 SKNAVITLRHGRCLCFMLRVGDLKRSMLISATIHQVVRKTTSPBGEVVLPHQVDIPENG 1800
QY 1801 VGGNGITFLVAPLIIYHYIDNSPLDYDLABSDLHHHDLLEIIVILGIVETTTGTTQARTS 1860
DB 1801 VGGNGITFLVAPLIIYHYIDNSPLDYDLABSDLHHHDLLEIIVILGIVETTTGTTQARTS 1860
QY 1861 VGGNGITFLVAPLIIYHYIDNSPLDYDLABSDLHHHDLLEIIVILGIVETTTGTTQARTS 1920
DB 1861 VGGNGITFLVAPLIIYHYIDNSPLDYDLABSDLHHHDLLEIIVILGIVETTTGTTQARTS 1920

QY 1861 YLADEILMGORFVPIVAEEDGRYSVDYSKFGNTIKVPTPLCTAROLDERS 1911
 DB 1841 YLADEILMGORFVPIVAEEDGRYSVDYSKFGNTIKVPTPLCTAROLDERS 1891

RESULT 5
 ID AAM57486 standard; protein, 1531 AA.
 AAM57486;
 AAM57486;
 14-AUG-1998 (first entry)
 Human MRP variant 1tPgpa (Lei/Pgpa).
 Multidrug resistance-associated protein, MRP, tumour, human, variant;
 multidrug resistance; MDR; leismania P-glycoprotein; 1tPgpa, Lei/Pgpa.
 Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 685 /label= L685S
 FT /note= "wild-type Leu is replaced by Ser"
 FT Misc-difference 1282 /label= R1282A
 FT /note= "wild-type Arg is replaced by Ala"
 FT
 XX
 XX US5766880-A.
 PN
 PD 16-JUN-1998.
 XX
 PF 05-JUN-1995; 95US-00463092.
 XX
 XX 27-OCT-1992; 92US-00966923.
 PR 08-MAR-1993; 93US-00029340.
 PR 26-OCT-1993; 93US-00141893.
 PR 20-MAR-1995; 95US-00407207.
 XX
 XX (TOOH) UNIV QUEBENS KINGSTON.
 PA
 XX
 PI Cole SP, Deeley RG;
 XX
 DR WPI; 1998-361687/31.
 DR N-PSDB; AAV31498.
 XX
 PT DNA encoding protein associated with multi-drug resistance - useful for
 PT as probe for identifying multi-drug resistant tumour cells.
 XX
 PS Claim 1; Col 67-78; 82pp; English.
 XX
 CC This represents a variant of the human multidrug resistance-associated
 CC protein (MRP). This natural variant is a leismania P-glycoprotein related
 CC molecule 1tPgpa (Lei/Pgpa). The human and murine MRP nucleic acid
 CC molecules can be used as probes for identifying multidrug resistant
 CC tumour cells by hybridisation to mRNA from tumour cells. The antisense
 CC nucleic acid can be used to reverse multidrug resistance (MDR). A
 CC recombinant expression vector containing the MRP nucleic acid molecules
 CC operatively linked to at least one regulatory sequence can be used to
 CC transform a host cell to produce a recombinant MDR-associated protein
 XX
 SQ Sequence 1531 AA;

Query Match 78.5%; Score 7860; DB 2; Length 1531;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGPSAGSPBIMDMNTNTSNDPTKCFONTLVVWPCYILMACFPFYLISRH 60
 DB 1 MALRGFCSAGSPBIMDMNTNTSNDPTKCFONTLVVWPCYILMACFPFYLISRH 60
 QY 61 DRGYIQWTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPFLIGITTLA 120

DB 61 DRGYIQWTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPFLIGITTLA 120
 QY 121 TPLIQERRRRGVSSGIMLTFWVALVCAIALRSKTMALKEDAQVDLPDITFYVFS 180
 DB 121 TPLIQERRRRGVSSGIMLTFWVALVCAIALRSKTMALKEDAQVDLPDITFYVFS 180
 QY 181 LLLIQVLSCPSRSPFSEITHDNPCESSASFSLRTFMWTTGLIVGYRPLGSD 240
 DB 181 LLLIQVLSCPSRSPFSEITHDNPCESSASFSLRTFMWTTGLIVGYRPLGSD 240
 QY 241 LMSLNKEDTSEQVVPVVKWKKCECAKTRQOPVNVSSKDPAPQKSSKVDANEVEAL 300
 DB 241 LMSLNKEDTSEQVVPVVKWKKCECAKTRQOPVNVSSKDPAPQKSSKVDANEVEAL 300
 QY 301 IVKSPQKSNPISLFKVLKTFPGPYFLMSFPFKAIHDMFSGPQILKLIFVNDTAPD 360
 DB 301 IVKSPQKSNPISLFKVLKTFPGPYFLMSFPFKAIHDMFSGPQILKLIFVNDTAPD 360
 QY 361 MOGYFTVLLFVTAQTLVTHQVPHICFVSGMKIKTAIVGAVRRKLVITNSARKSTV 420
 DB 361 MOGYFTVLLFVTAQTLVTHQVPHICFVSGMKIKTAIVGAVRRKLVITNSARKSTV 420
 QY 421 GEIVNLMSVDAORFMDLATYINMISAPLQVIALYLLMNLGSPVLAGVAVMVLAMYVN 480
 DB 421 GEIVNLMSVDAORFMDLATYINMISAPLQVIALYLLMNLGSPVLAGVAVMVLAMYVN 480
 QY 481 AVAMKTKTYQVAMHMSKDNRIKLMEIINGIKVLKYLAMELAKDYLAIRBELKVLK 540
 DB 481 AVAMKTKTYQVAMHMSKDNRIKLMEIINGIKVLKYLAMELAKDYLAIRBELKVLK 540
 QY 541 KSAVLSAVGTFVTCVCFELVALCTFAVYVITIDENIIDAQTAFLSALFNILRPLNLP 600
 DB 541 KSAVLSAVGTFVTCVCFELVALCTFAVYVITIDENIIDAQTAFLSALFNILRPLNLP 600
 QY 601 MWISSIVQASVLSKRLRIFLSHELEBPSJERRPVKDGGSNTSYVNAATFTMARSDPT 660
 DB 601 MWISSIVQASVLSKRLRIFLSHELEBPSJERRPVKDGGSNTSYVNAATFTMARSDPT 660
 QY 661 LMGITFSIPGALVAVVGQVCGKSSLSLALMEQVBEHVAIKGSVAVPOQAWIQND 720
 DB 661 LMGITFSIPGALVAVVGQVCGKSSLSLALMEQVBEHVAIKGSVAVPOQAWIQND 720
 QY 721 SLRENTIFGQLEBPYRSVIOACALLPDLIELPSGDRTEIGKGVNLSCGOKRVSILAR 780
 DB 721 SLRENTIFGQLEBPYRSVIOACALLPDLIELPSGDRTEIGKGVNLSCGOKRVSILAR 780
 QY 781 AVYSNADIVLFDPLSLAVDAHVGHKIFENVIGKMKLNKTRILVTHSMGYLFQVDVITV 840
 DB 781 AVYSNADIVLFDPLSLAVDAHVGHKIFENVIGKMKLNKTRILVTHSMGYLFQVDVITV 840
 QY 841 MSGGKISEMGSYQELLARDGAFELRTYASTEOBDAENGVGVSGPGEAKKOMENGM 900
 DB 841 MSGGKISEMGSYQELLARDGAFELRTYASTEOBDAENGVGVSGPGEAKKOMENGM 900
 QY 901 LYTDSAGKQLOROLSSSSSYSGDISRHNSTALQRAEAKKETWKLMEADKQOTGVKL 960
 DB 901 LYTDSAGKQLOROLSSSSSYSGDISRHNSTALQRAEAKKETWKLMEADKQOTGVKL 960
 QY 961 SYVMDYMKIIGLISFISFLFMCNHYVSAASNYWLSIMTDDIIVNGTOHSTVRLSVYG 1020
 DB 961 SYVMDYMKIIGLISFISFLFMCNHYVSAASNYWLSIMTDDIIVNGTOHSTVRLSVYG 1020
 QY 1021 ALGISOGIAVFGYSMAVISIGILIASRCLVLDLHSLRSPMSFERTPSGNLVNRSKEL 1080
 DB 1021 ALGISOGIAVFGYSMAVISIGILIASRCLVLDLHSLRSPMSFERTPSGNLVNRSKEL 1080
 QY 1081 DTVDSMIPVYIKMFMGSLPNTVACVILLATPILAIIPPLGLIYFVVRPFVASSROL 1140
 DB 1081 DTVDSMIPVYIKMFMGSLPNTVACVILLATPILAIIPPLGLIYFVVRPFVASSROL 1140
 QY 1141 KRLESVRSFVYSHFNETLLGVSVIRAFEEQEFIIQSDDLKVDENKAYVPSIVANRWLA 1200

Db 1141 KRLESVSRSPVYSHFNETLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPYSIVANRWLA 1200
QY 1201 VRLECVNGCIYLFALPAFVAVSRHSLSAGLVGLSVSYLQVTTYINMLVRMSSEMETNIVA 1260
Db 1201 VRLECVNGCIYLFALPAFVAVSRHSLSAGLVGLSVSYLQVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIQETAPPSWQVGRVEFRNYCLARYREDLDFLRIHINVTINGG 1330
Db 1261 VERLKEYSETEKEAPWQIQETAPPSWQVGRVEFRNYCLARYREDLDFLRIHINVTINGG 1330
QY 1321 EKVGIVERTGAGKSLTGLFRINESAGEIITIDGINAKIGHDLRKITITIDPVLTF 1380
Db 1321 EKVGIVERTGAGKSLTGLFRINESAGEIITIDGINAKIGHDLRKITITIDPVLTF 1380
QY 1381 SGSLRMNLDPFSGYSDSEWMTSLBLAHLDKDFVSALPDLDHECAGGENTLSVGOROLVCL 1440
Db 1381 SGSLRMNLDPFSGYSDSEWMTSLBLAHLDKDFVSALPDLDHECAGGENTLSVGOROLVCL 1440
QY 1441 APALLRKTIIVLDEBATAVDLETFDDLIQSTIRFQEDCTVLTAAHRLNTIMDYTRIVL 1500
Db 1441 APALLRKTIIVLDEBATAVDLETFDDLIQSTIRFQEDCTVLTAAHRLNTIMDYTRIVL 1500
QY 1501 DKGEIOEYGAPSDLDLQORGLFYMAKXGALV 1531
Db 1501 DKGEIOEYGAPSDLDLQORGLFYMAKXGALV 1531

RESULT 6

AAM74471

ID AAM74471 standard; protein; 1531 AA.

AC AAM74471;

DT 18-MAY-1999 (first entry)

DE Human multidrug resistance-associated protein variant.

KM Multidrug resistance-associated protein; MDR; human; diagnosis;

KW MDR tumour cell identification; cancer therapy.

OS Homo sapiens.

PN US5882875-A.

PD 16-MAR-1999.

PE 05-JUN-1995; 95US-00462109.

PR 27-OCT-1992; 92US-00966923.

PR 08-MAR-1993; 93US-00029340.

PR 26-OCT-1993; 93US-00141893.

PR 20-MAR-1995; 95US-00407207.

PA (TOOH) UNIV QUEBENS KINGSTON.

PI Cole SPC, Deeley RG;

DR WPI: 1999-214061/18.

PT N-PSDB; AAX21977.

PT Identifying a multidrug resistant tumor cell by contacting the cell with

PT an antibody/antigen-binding fragment - which binds to an expressed

PT protein encoded by multidrug resistance-associated protein (MRP) nucleic

PT acid.

PS Claim 3; Col 69-80; 80pp; English.

XX This sequence is the human multidrug resistance-associated (MDR) protein.

XX The invention relates to a method for identifying a multidrug resistant

XX (MDR) tumour cell. Compositions and methods utilising the MDR proteins

XX can be used to treat patients with tumours displaying multidrug

XX resistance, particularly those displaying resistance to anthracyclines,

XX epipodophyllotoxins, vinca alkaloids, and hydrophobic drugs. The methods

CC for inhibiting/killing a MDR tumour cell can be useful for treating
CC breast cancer, leukaemias, fibrosarcomas, cervical cancer, gliomas,
CC thymomas, neuroblastomas and lung cancer. The MDR DNA sequences when
CC labeled are useful as molecular probes for diagnosing multidrug
CC resistance of a tumour (using cells from a tumour biopsy) and for
CC designing ribozymes which are capable of cleaving a single-stranded
CC nucleic acid encoding a protein having MRP activity. Recombinant
CC expression vectors containing human MDR coding sequences can be
CC transfected into a drug sensitive cell line to produce a protein in the
CC cell which confers MDR, protecting non-resistant non-tumour cells from
CC the effects of chemotherapeutic agents has major clinical importance. Cells
CC transformed with the MDR coding sequences are useful for testing
CC potential therapeutic agents for their effectiveness against MDR cells
CC and for identifying chemosensitizers of a therapeutic agent
XX
SQ Sequence 1531 AA;

Query Match 78.5%; Score 7860; DB 2; Length 1531;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Gaps 0;

Matches 1531; Conservative 0; Indels 0; Gaps 0;

QY 1 MALRGFCSADGSDPLMDMNTVMTNSNDPTKCFQNTVIVWPCCPYLMACFPFYFLYSRH 60
Db 1 MALRGFCSADGSDPLMDMNTVMTNSNDPTKCFQNTVIVWPCCPYLMACFPFYFLYSRH 60
QY 61 DRGYIQMTPLNKTKTALGFLIMIVCMADLFYSFWSRSGIFLAEVFLVSPITLIGITLLA 120
Db 61 DRGYIQMTPLNKTKTALGFLIMIVCMADLFYSFWSRSGIFLAEVFLVSPITLIGITLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFMVLVYALCALILRSKIMTALKENAOVDLPDITFYFYFS 180
Db 121 TFLIQLERRKGVSSGIMLTFMVLVYALCALILRSKIMTALKENAOVDLPDITFYFYFS 180
QY 181 LLILQLVLSGFSDSPLFSETIHDNCPSPSSASFLSRTIFPMWITGLIVRGYROPLEGSD 240
Db 181 LLILQLVLSGFSDSPLFSETIHDNCPSPSSASFLSRTIFPMWITGLIVRGYROPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSXDPAPKSSKVDANEVEAL 300
Db 241 LMSLNKEDTSEQVVPVLVKNWKECAKTRKQPVKVYSSXDPAPKSSKVDANEVEAL 300
QY 301 IVKSPQKEMNSLKFVLYKTGPGYFLMSFPFKAIHDLMSGPDILKLLIFVNDYRAPD 360
Db 301 IVKSPQKEMNSLKFVLYKTGPGYFLMSFPFKAIHDLMSGPDILKLLIFVNDYRAPD 360
QY 361 WQGFYTVLLFVTACLOTLVHGYEHI CFVSGMAIKTAVGAVYRKALVITNSARKSTV 420
Db 361 WQGFYTVLLFVTACLOTLVHGYEHI CFVSGMAIKTAVGAVYRKALVITNSARKSTV 420
QY 421 GEIVNLMSVDAQREMDLATYINMTWSAPLQVTLALYLLMNLGSPVLGAVAVMYLMPVN 480
Db 421 GEIVNLMSVDAQREMDLATYINMTWSAPLQVTLALYLLMNLGSPVLGAVAVMYLMPVN 480
QY 481 AVMAKTKTYOVAVAMSKDNRIKLMNELLNGIKYLYAMELAKDKYLAIRQSELYLK 540
Db 481 AVMAKTKTYOVAVAMSKDNRIKLMNELLNGIKYLYAMELAKDKYLAIRQSELYLK 540
QY 541 KSAVLAVGFTWCTPFLVALCTFAVYVITDENNNIIIDAOFAFSLAFNLRPEPLNLP 600
Db 541 KSAVLAVGFTWCTPFLVALCTFAVYVITDENNNIIIDAOFAFSLAFNLRPEPLNLP 600
QY 601 MVISSIVQASVSLRLRILFISHELEPDSIERRPVKDGGTNSITVENATPTWARSDEPT 660
Db 601 MVISSIVQASVSLRLRILFISHELEPDSIERRPVKDGGTNSITVENATPTWARSDEPT 660
QY 661 LNGITFSPREGALVAVVGVQCGKSSLLSALLAEMDKVEGHVAKGSVAVYVPOQAMIOND 720
Db 661 LNGITFSPREGALVAVVGVQCGKSSLLSALLAEMDKVEGHVAKGSVAVYVPOQAMIOND 720
QY 721 SLRENIILFGCLLEPPYRSVYQACALLPDLEILSSGRTEIGEGVNLSSGOKORVSLAR 780
Db 721 SLRENIILFGCLLEPPYRSVYQACALLPDLEILSSGRTEIGEGVNLSSGOKORVSLAR 780

QY 781 AVYSNADLYLDEDDLSAVDAHVGHIFENVIIPGKMLKNKTRILVTHSMSTLPQVDYIV 840
 DB 781 AVYSNADLYLDEDDLSAVDAHVGHIFENVIIPGKMLKNKTRILVTHSMSTLPQVDYIV 840
 QY 841 MSGGKISEMGSYQOELARDGAFAFLTYASTEOEODAEENGVTGSGPKKAYOMENGM 900
 DB 841 MSGGKISEMGSYQOELARDGAFAFLTYASTEOEODAEENGVTGSGPKKAYOMENGM 900
 QY 901 LVYDSACKOLOROLSSSSSSSGDISRRHNSTAEIOKAEKKEETWKLMEADKATQGVKL 960
 DB 901 LVYDSACKOLOROLSSSSSSSGDISRRHNSTAEIOKAEKKEETWKLMEADKATQGVKL 960
 QY 961 SVYWDYKKAIGLFIISFLIFLFCNHYSAALSNYMLSLMTDPIVNGTOBHTKRLSYVG 1020
 DB 961 SVYWDYKKAIGLFIISFLIFLFCNHYSAALSNYMLSLMTDPIVNGTOBHTKRLSYVG 1020
 QY 1021 ALGISOGIAVGVSMASVIGGILASRCLHVDLHSLIRSPMSFEPTPSGNLVNRFSEKL 1080
 DB 1021 ALGISOGIAVGVSMASVIGGILASRCLHVDLHSLIRSPMSFEPTPSGNLVNRFSEKL 1080
 QY 1081 DTWDSMTPEYIKMFMGSLFNVIIGACIYILATPIAIIIPPLGLIYFFVQRFYASSROL 1140
 DB 1081 DTWDSMTPEYIKMFMGSLFNVIIGACIYILATPIAIIIPPLGLIYFFVQRFYASSROL 1140
 QY 1141 KRLESVSRSPVYSHFNFTLGVSVYIRAFEBEORFIHOSDLKVDENOKAYPSIVANRMLA 1200
 DB 1141 KRLESVSRSPVYSHFNFTLGVSVYIRAFEBEORFIHOSDLKVDENOKAYPSIVANRMLA 1200
 QY 1201 VRLECVCNCIYLFALFAVIRSRHSLASGLVGLSVSYSLQVTTYLVNMLVRMSSEMETIVA 1260
 DB 1201 VRLECVCNCIYLFALFAVIRSRHSLASGLVGLSVSYSLQVTTYLVNMLVRMSSEMETIVA 1260
 QY 1261 VERLKESETEKEKPMQIOETAPSSWPQVGRVFNRYCIRYREDLDFVLRHINVTNGG 1320
 DB 1261 VERLKESETEKEKPMQIOETAPSSWPQVGRVFNRYCIRYREDLDFVLRHINVTNGG 1320
 QY 1321 EKYGIIVGRTGAGKSLTLGLFRINESAGEIIDIINIAKIGLHDLFKITIIIPQDVLVF 1380
 DB 1321 EKYGIIVGRTGAGKSLTLGLFRINESAGEIIDIINIAKIGLHDLFKITIIIPQDVLVF 1380
 QY 1381 SSGSLRNLDPFSQYSDSEEVWTSLELAHLKDFVSALPKLDHECAEGEENL SVGQROLVCL 1440
 DB 1381 SSGSLRNLDPFSQYSDSEEVWTSLELAHLKDFVSALPKLDHECAEGEENL SVGQROLVCL 1440
 QY 1441 ARALLRKTKLIVDEATAVADLETFDULIOSTIRTOPEFDCVLTIAHRLNTIMDYTRIVL 1500
 DB 1441 ARALLRKTKLIVDEATAVADLETFDULIOSTIRTOPEFDCVLTIAHRLNTIMDYTRIVL 1500
 QY 1501 DKGEIOEYGAPSDLQORGLFYSMAKDAGLV 1531
 DB 1501 DKGEIOEYGAPSDLQORGLFYSMAKDAGLV 1531

XX 27-OCT-1992; 92US-00966923.
 PR 08-MAR-1993; 93US-00029340.
 PR 26-OCT-1993; 93US-00141893.
 PR 20-MAR-1995; 95US-00407207.
 PA (TOOH) UNIV QUEBENS KINGSTON.
 XX
 PI Cole SPC, Deeley RG;
 XX
 DR WPI; 1999-253868/21.
 DR N-PSDB; AAX19818.
 XX
 PT Protecting mammalian cells against cytotoxic drugs.
 XX
 PS Claim 2; Col 79-86; 82pp; English.
 CC
 CC The present sequence represents a human multidrug resistance-associated
 CC protein (MRP). The present invention also describes a method for
 CC protecting a mammalian cell against the cytotoxicity of anthracyclines,
 CC epipodophyllotoxins and Vinca alkaloids (A) by introducing into it a
 CC nucleic acid (I) that hybridizes under stringent conditions to a nucleic
 CC acid (II) that encodes an MRP protein (III). Introduction of (I) protects
 CC cells against cytotoxic effects of (A), particularly to protect normal
 CC cells against (A) being used for treatment of cancers. Cells transformed
 CC with (I) can be used to screen for agents that affect multidrug
 CC resistance or are directly toxic to multidrug resistant cells, i.e.
 CC potential therapeutics for multidrug-resistant cancers. Confering
 CC resistance to normal cells should allow an increase in the dose of (A)
 CC that can be administered safely
 XX
 SQ Sequence 1531 AA:
 Query Match 78.5%; Score 7860; DB 2; Length 1531;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MALRGFSGADSGDPLMDMNTWNTSNPDFKCKQNTLVVWPCFYLMACPPFFLYLSRH 60
 DB 1 MALRGFSGADSGDPLMDMNTWNTSNPDFKCKQNTLVVWPCFYLMACPPFFLYLSRH 60
 QY 61 DRGYIOWTPLNKTKTALGFLMTVCNADLYFSMERSRGFLAPVFLVSTLIGITTLA 120
 DB 61 DRGYIOWTPLNKTKTALGFLMTVCNADLYFSMERSRGFLAPVFLVSTLIGITTLA 120
 QY 121 TFLIOERRKGVSSGIMLTFMLVALVCAIAIRSKIMTALKEADAQVDLFRDITFYVFS 180
 DB 121 TFLIOERRKGVSSGIMLTFMLVALVCAIAIRSKIMTALKEADAQVDLFRDITFYVFS 180
 QY 181 LLLIOVLSCFSRSPPLFSETTHDNPCESSASFLSRITFWITGLIYVGRQPLBGS 240
 DB 181 LLLIOVLSCFSRSPPLFSETTHDNPCESSASFLSRITFWITGLIYVGRQPLBGS 240
 QY 241 LWSLNKEDPSEGVVPLVKNWKECAKTRQPKVYSSDPOPKSSSVVDNNEVEAL 300
 DB 241 LWSLNKEDPSEGVVPLVKNWKECAKTRQPKVYSSDPOPKSSSVVDNNEVEAL 300
 QY 301 IVKSPQKWNPSLFLKVLKTFGPFYLSMFEFFKAIDHLMFSGQIILKLIKFNNDTKAPD 360
 DB 301 IVKSPQKWNPSLFLKVLKTFGPFYLSMFEFFKAIDHLMFSGQIILKLIKFNNDTKAPD 360
 QY 361 WQGYFTVLLFFTACIQTLVLAHOYFHI CPVSGMRITAVIGAYRRALVITNSARKSSTV 420
 DB 361 WQGYFTVLLFFTACIQTLVLAHOYFHI CPVSGMRITAVIGAYRRALVITNSARKSSTV 420
 QY 421 GEIVNLMSYDAORFMDLATYINNIWSAPLOVITALLVLMNLGSPSYLAGAVWVLPVN 480
 DB 421 GEIVNLMSYDAORFMDLATYINNIWSAPLOVITALLVLMNLGSPSYLAGAVWVLPVN 480
 QY 481 AVAMKTKTYQVAHMSKONRIKLANEILNGIKVLKYAMELAFKDKVLAIRQEBLKVLK 540
 DB 481 AVAMKTKTYQVAHMSKONRIKLANEILNGIKVLKYAMELAFKDKVLAIRQEBLKVLK 540

```
QY 541 KSAVLASVGFETFWCTPFLVALCTPAVYVITDENNIIDAQTAFAVSLAFNILEPPLILP 600
DB 541 KSAVLASVGFETFWCTPFLVALCTPAVYVITDENNIIDAQTAFAVSLAFNILEPPLILP 600
QY 601 MVISSIVQASVSLKRLRIPLSHELEPDSIRRPVKOGGNGNSTVNAATATMARSDPPT 660
DB 601 MVISSIVQASVSLKRLRIPLSHELEPDSIRRPVKOGGNGNSTVNAATATMARSDPPT 660
QY 661 LINGITFSPREGALVAVVQVCGCKSSLLSALLAEMDKYEGVAIKGSVAVYPOQAMIOND 720
DB 661 LINGITFSPREGALVAVVQVCGCKSSLLSALLAEMDKYEGVAIKGSVAVYPOQAMIOND 720
QY 721 SLRENILFGCQLEBPYRSVIOACALLPDLEILPSGDRTEIGEKVNLISGGQXORSILAR 780
DB 721 SLRENILFGCQLEBPYRSVIOACALLPDLEILPSGDRTEIGEKVNLISGGQXORSILAR 780
QY 781 AVYNNADIYLFDDPLSANDAHVGHIFENVIQPKGMLKNKRIILVTHSMSTLPQVDVYIV 840
DB 781 AVYNNADIYLFDDPLSANDAHVGHIFENVIQPKGMLKNKRIILVTHSMSTLPQVDVYIV 840
QY 841 MSGGKISEMGSYOELLARDGAFAEFLRTYASTEODAEENGVTGVSQPKGKAKOMENGM 900
DB 841 MSGGKISEMGSYOELLARDGAFAEFLRTYASTEODAEENGVTGVSQPKGKAKOMENGM 900
QY 901 LVTDASAGKOLQROUSSSSSSYSGDISRHHNSTAELOKAKAKEETWKLMEADKAQTGYKL 960
DB 901 LVTDASAGKOLQROUSSSSSSYSGDISRHHNSTAELOKAKAKEETWKLMEADKAQTGYKL 960
QY 961 SVYMDYMKALIFLSLISLIFMCHNVASLANSWLSIWTDDPIVNGOETHTKRLSYYG 1020
DB 961 SVYMDYMKALIFLSLISLIFMCHNVASLANSWLSIWTDDPIVNGOETHTKRLSYYG 1020
QY 1021 ALGISOGIAVFGYSMAVSIIGGILASRCLHVDLHSHILSPMSPFERTSGNIVNRFSEL 1080
DB 1021 ALGISOGIAVFGYSMAVSIIGGILASRCLHVDLHSHILSPMSPFERTSGNIVNRFSEL 1080
QY 1081 DTVDMSIPEVIKMFMSLFENVIGACTIVILLATPIAIIIPGLGIYFFVGFRYYASSRQL 1140
DB 1081 DTVDMSIPEVIKMFMSLFENVIGACTIVILLATPIAIIIPGLGIYFFVGFRYYASSRQL 1140
QY 1141 KRLSVSVSPYVSHPHNETLIGSVIVIRAFEBQERFHQSDLVADENQKAYYSIVANRWLA 1200
DB 1141 KRLSVSVSPYVSHPHNETLIGSVIVIRAFEBQERFHQSDLVADENQKAYYSIVANRWLA 1200
QY 1201 VRLECVGNCIYLFALPAFVISRHSLSAGLVGSLVSYSIQVTTYLAWLVRMSSEMETNIVA 1260
DB 1201 VRLECVGNCIYLFALPAFVISRHSLSAGLVGSLVSYSIQVTTYLAWLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPWQIOETAPPSWPQVGRVEFRNYCLARYEDLDFVLRIHINTINGG 1320
DB 1261 VERLKEYSETEKEAPWQIOETAPPSWPQVGRVEFRNYCLARYEDLDFVLRIHINTINGG 1320
QY 1321 EKVGIIVGRTGAGKSLTIGLFRINSAAGEIIIDGINIAKIGLHLRKRITIIIPDDPYLF 1380
DB 1321 EKVGIIVGRTGAGKSLTIGLFRINSAAGEIIIDGINIAKIGLHLRKRITIIIPDDPYLF 1380
QY 1381 SGLSRMNLDPSPSOVSDEEWMTSLELAHLKDFYSALPDKLDEHCAGGENLSVGORQIYCL 1440
DB 1381 SGLSRMNLDPSPSOVSDEEWMTSLELAHLKDFYSALPDKLDEHCAGGENLSVGORQIYCL 1440
QY 1441 ARALLRKTILVLEBATAVAVDLETDLLIOSTIRTOFEDCTVLLIARLNTIMDYRYIVL 1500
DB 1441 ARALLRKTILVLEBATAVAVDLETDLLIOSTIRTOFEDCTVLLIARLNTIMDYRYIVL 1500
QY 1501 DKGEIOEYGAPSDLLQORGLFYSMKADGLV 1531
DB 1501 DKGEIOEYGAPSDLLQORGLFYSMKADGLV 1531
```

```
RESULT 8
AAVS5799
ID AAVS5799 standard; protein; 1531 AA.
XX
```

```
AC AAVS5799;
XX
XX 28-FEB-2000 (first entry)
XX
DE Human multidrug resistance-associated protein (MRP) variant.
XX
XX Chemosensitizer; multidrug resistance-associated protein; MRP; human;
XX therapeutic agent; P-glycoprotein-mediated multidrug resistance; lung;
XX cancer; variant.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Key 685
XX Misc-difference /label= L685S
XX /note= "wild-type Leu is replaced by Ser"
XX
XX Misc-difference 1282
XX /label= R1282A
XX /note= "wild-type Arg is replaced by Ala"
XX
XX US6001563-A.
XX
XX PD 14-DEC-1999.
XX
XX PF 05-JUN-1995; 95US-00463179.
XX
XX PR 27-OCT-1992; 92US-00966923.
XX PR 08-MAR-1993; 93US-00029340.
XX PR 26-OCT-1993; 93US-00141893.
XX PR 20-MAR-1995; 95US-00407207.
XX
XX PA (TOOH ) UNIV QUEBENS KINGSTON.
XX
XX PI Cole SP, Deeley RG;
XX
XX WP1; 2000-061877/05.
XX
XX DR N-PSDB; AAZ39556.
XX
XX PT Identification of chemosensitizers useful for treating cancer, using
XX nucleic acid encoding multidrug resistance-associated protein.
XX
XX PS Claim 3; Col 69-80; 77bp; English.
XX
XX The invention provides a method for identifying a substance which is a
XX chemosensitizer that comprises, contacting a cell transfected with
XX nucleic acid encoding multidrug resistance-associated protein (MRP) with
XX a therapeutic agent in vitro. The method is useful for identifying
XX chemosensitizers which may then be used to treat cancer (especially lung
XX cancer). The method allows the identification of chemosensitizers which
XX do not reverse P-glycoprotein-mediated multidrug resistance. The present
XX sequence represents a human MRP variant
XX
XX Sequence 1531 AA:
XX
XX Query Match 78.5%; Score 7860; DB 3; Length 1531;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MALHFGSADSDPLMNMNTVMTNSNDPFTKCFONTYLVWPCGYIWAACPFYFLYLSRH 60
DB 1 MALHFGSADSDPLMNMNTVMTNSNDPFTKCFONTYLVWPCGYIWAACPFYFLYLSRH 60
QY 61 DRGYIOWTPLNKKTALGFLLMIYCMADLFYSFWERSRGJFLAVVFLVSPFLIGITLLA 120
DB 61 DRGYIOWTPLNKKTALGFLLMIYCMADLFYSFWERSRGJFLAVVFLVSPFLIGITLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFWLVALVCALALRSKIMTALKEDAQVDLFFDITFYVYFS 180
DB 121 TFLIQLERRKGVSSGIMLTFWLVALVCALALRSKIMTALKEDAQVDLFFDITFYVYFS 180
QY 181 LLLIQLVLSGFSRSPFLFSETIHDPNPCPSSASFLSRITRWMTTGLIVRGYRQPLRGSD 240
DB 181 LLLIQLVLSGFSRSPFLFSETIHDPNPCPSSASFLSRITRWMTTGLIVRGYRQPLRGSD 240
```

QY 241 LMSLNKEDTSEOVVPLVKWKKCECAKTRKOPVKVYSSKDPAPQKSSSKVDANEVEAL 300
 DB 241 LMSLNKEDTSEOVVPLVKWKKCECAKTRKOPVKVYSSKDPAPQKSSSKVDANEVEAL 300
 QY 301 IVKSPQKEMNPDLFEKVLKFTGPGFLMSFFPKAIDHLMFSGPOLIKL1KFVNDTRAPD 360
 DB 301 IVKSPQKEMNPDLFEKVLKFTGPGFLMSFFPKAIDHLMFSGPOLIKL1KFVNDTRAPD 360
 QY 361 WQGFYVVLVFNTRCLOTLVHQYFHI CFVSGMFKTNAVIGAVYRKALVITNSARKSTV 420
 DB 361 WQGFYVVLVFNTRCLOTLVHQYFHI CFVSGMFKTNAVIGAVYRKALVITNSARKSTV 420
 QY 421 GEIYNLMSVDAOREFMDLATYINM1WSAPLOVITLALYLLMLNGPSYAGVAVWVLAMPVN 480
 DB 421 GEIYNLMSVDAOREFMDLATYINM1WSAPLOVITLALYLLMLNGPSYAGVAVWVLAMPVN 480
 QY 481 AVNMAKTKTYOVAHMKSKDNRIKLMNEILNGIKVLYAMELAFKDKVLA1ROBELKVLK 540
 DB 481 AVNMAKTKTYOVAHMKSKDNRIKLMNEILNGIKVLYAMELAFKDKVLA1ROBELKVLK 540
 QY 541 KSAVLASVGTFTWCTPFLVALCTPFAVYVITDENNIIDAQTAFSVSLFNLIRPPLILP 600
 DB 541 KSAVLASVGTFTWCTPFLVALCTPFAVYVITDENNIIDAQTAFSVSLFNLIRPPLILP 600
 QY 601 MVISSIVQASVSLKRLIPLSHHELEPDSTERRPVKGGGNTITVNAFTMARSDPT 660
 DB 601 MVISSIVQASVSLKRLIPLSHHELEPDSTERRPVKGGGNTITVNAFTMARSDPT 660
 QY 661 LNGITFSIPGALVAVVGVCCKGKSLSLALLAEMDKYEGHVAIKGSVAVYVPOQAW1OND 720
 DB 661 LNGITFSIPGALVAVVGVCCKGKSLSLALLAEMDKYEGHVAIKGSVAVYVPOQAW1OND 720
 QY 721 SLRENIILFGCOLEBPYRSYIQAACALLPDLLEILPDSGRTEIGEKNVLSGGQKQVSLAR 780
 DB 721 SLRENIILFGCOLEBPYRSYIQAACALLPDLLEILPDSGRTEIGEKNVLSGGQKQVSLAR 780
 QY 781 AVVSNADIYLPDPLSAVDAHVGHIFENYIGPQGMKKNTRILLVTHSMSTLPDVAIIV 840
 DB 781 AVVSNADIYLPDPLSAVDAHVGHIFENYIGPQGMKKNTRILLVTHSMSTLPDVAIIV 840
 QY 841 MSGGKISEMGSYOELIARDAFAFLRTYASTEOQDAENGVTGSGPGKEAKOMENGM 900
 DB 841 MSGGKISEMGSYOELIARDAFAFLRTYASTEOQDAENGVTGSGPGKEAKOMENGM 900
 QY 901 LVTDGAKQIQRULSSSSSYSGDISRHNSSTAEIQAKEAKEETWKLMEADKAQTQVKL 960
 DB 901 LVTDGAKQIQRULSSSSSYSGDISRHNSSTAEIQAKEAKEETWKLMEADKAQTQVKL 960
 QY 961 SYVWDVWKATIGLFTSEFLFLPMCNHVSALASNWLSLMTDDPIVNGTOHTKRLSVYG 1020
 DB 961 SYVWDVWKATIGLFTSEFLFLPMCNHVSALASNWLSLMTDDPIVNGTOHTKRLSVYG 1020
 QY 1021 ALGISOGIAVFGXSMVAISGIIASRCLAHYDLHSIRSPMSPFERTPSGNLVNRSKEL 1080
 DB 1021 ALGISOGIAVFGXSMVAISGIIASRCLAHYDLHSIRSPMSPFERTPSGNLVNRSKEL 1080
 QY 1081 DTVDSMIPBVIKMFMSGLFNVIAGACIVILLATP1AIIIPPLGLIYFVORFVYASSROL 1140
 DB 1081 DTVDSMIPBVIKMFMSGLFNVIAGACIVILLATP1AIIIPPLGLIYFVORFVYASSROL 1140
 QY 1141 KRLESVRSRPFVYSHFNFTLLGVSVYTRAFEOERFIHOSDLKVDENQAAVYPSIYANRWLA 1200
 DB 1141 KRLESVRSRPFVYSHFNFTLLGVSVYTRAFEOERFIHOSDLKVDENQAAVYPSIYANRWLA 1200
 QY 1201 VRLCEVGNCLVLPALFAVISRHSLSAGVLGSYSGLQVTTYLNLVNRMSSEMETIYA 1260
 DB 1201 VRLCEVGNCLVLPALFAVISRHSLSAGVLGSYSGLQVTTYLNLVNRMSSEMETIYA 1260
 QY 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVEFRNYCLRYREDLDFVLRIHNVITNGG 1320
 DB 1261 VERLKEYSETEKAPWQIOETAPSSWPQVGRVEFRNYCLRYREDLDFVLRIHNVITNGG 1320

QY 1321 EKVGIIVRTGAGKSLTLGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPLV 1380
 DB 1321 EKVGIIVRTGAGKSLTLGLFRINESAGEEIIIDGINIAKIGLHDLRFKTIIPQDPLV 1380
 QY 1381 SCSLRNMLDPFSQYSDEEVTSLLELAHLKDFVSALPDKLDHECAEGGENLSVCGROLVCL 1440
 DB 1381 SCSLRNMLDPFSQYSDEEVTSLLELAHLKDFVSALPDKLDHECAEGGENLSVCGROLVCL 1440
 QY 1441 ARALLRRTKTLVDEATAAADLTETDLOSTITRQPEDCVTLTAHRLNTIMYTRIVL 1500
 DB 1441 ARALLRRTKTLVDEATAAADLTETDLOSTITRQPEDCVTLTAHRLNTIMYTRIVL 1500
 QY 1501 DKGEIOEYGAAPSDLLQORGLFYSMAKADAGLV 1531
 DB 1501 DKGEIOEYGAAPSDLLQORGLFYSMAKADAGLV 1531

RESULT 9
 AAY78873
 ID AAY78873 standard; protein; 1531 AA.
 XX AAY78873;
 AC AAY78873;
 XX 19-MAY-2000 (first entry)
 DT 19-MAY-2000 (first entry)
 XX
 DE Multidrug resistance protein (MRP) natural variant amino acid sequence.
 XX
 KW Multidrug resistance protein; MRP; human; anthracycline; vinca alkaloid;
 KM epipodophyllotoxin; cancer; leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN US6025473-A.
 XX 15-FEB-2000.
 PD 15-FEB-2000.
 XX
 PF 05-JUN-1995; 95US-00461384.
 XX
 XX 27-OCT-1992; 92US-00966923.
 PR 08-MAR-1993; 93US-00029340.
 PR 26-OCT-1993; 93US-00141893.
 PR 20-MAR-1995; 95US-00407207.
 XX
 PA (TOOH) UNIT QUBENS KINGSTON.
 XX
 PI Cole SPC, Deeley RG;
 XX
 DR WPI; 2000-181838/16.
 DR N-PDB; AA290193.
 XX
 XX Isolated protein conferring multidrug resistance, to at least two drugs
 PT selected from anthracyclines, epipodophyllotoxins and vinca alkaloids, on
 PT a drug sensitive mammalian cell.
 XX
 PS Claim 10; Col 79-88; 78pp; English.
 XX
 CC This sequence represents a human multidrug resistance protein (MRP)
 CC natural variant amino acid sequence. The human MRP confers multidrug
 CC resistance, including resistance to at least two drugs selected from
 CC anthracyclines, epipodophyllotoxins and vinca alkaloids, on a drug
 CC sensitive mammalian cell, when the protein is expressed in the cell. The
 CC multidrug resistance is not substantially reversed by chemosensitizers
 CC which reverse P-glycoprotein-mediated multidrug resistance. The MRP
 CC protein sequence can be used to generate antibodies against MRP. The MRP
 CC protein and nucleotide sequences can be used in compositions which are
 CC used to treat patients with tumours displaying multidrug resistance. The
 CC compositions and methods of the invention can be used particularly to
 CC treat breast cancer, leukaemia, fibrosarcomas, cervical cancer, gliomas,
 CC thymomas, neuroblastomas, and lung cancer. Antibodies directed against
 CC MRP can be used to inhibit the multidrug resistance of a multidrug
 CC resistant cell
 XX
 SO Sequence 1531 AA;

Query Match 78.5%; Score 7860; Db 3; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MALRGFSSAGSDPLMDMNTMTNTSNDPTKCFONTYLVWVPCFYLMACFPFYLYSRH 60
DB 1 MALRGFSSAGSDPLMDMNTMTNTSNDPTKCFONTYLVWVPCFYLMACFPFYLYSRH 60
QY 61 DRGVIQMTPLNKTALGFLIMICVADLFYSFMSRSGIFLAIVFLVSPILLGITLLA 120
DB 61 DRGVIQMTPLNKTALGFLIMICVADLFYSFMSRSGIFLAIVFLVSPILLGITLLA 120
QY 121 TFLIOLERRKGVSSGIMLFMWLVALCALILRSKIMTALKEDAOYDLPRDITFYFFYS 180
DB 121 TFLIOLERRKGVSSGIMLFMWLVALCALILRSKIMTALKEDAOYDLPRDITFYFFYS 180
QY 181 LLLIQVLVSCFSDSPFLFSETIHDNPNCPRESSASFSLRITFWITGLIVRGYRQPLGSD 240
DB 181 LLLIQVLVSCFSDSPFLFSETIHDNPNCPRESSASFSLRITFWITGLIVRGYRQPLGSD 240
QY 241 LMSLNKEDTSEQVAVLVKMKKCAKTRKQPVKVYVSSKDPAPKSSKYDANEVEAL 300
DB 241 LMSLNKEDTSEQVAVLVKMKKCAKTRKQPVKVYVSSKDPAPKSSKYDANEVEAL 300
QY 301 IVNSPQKEMNPSLFKVLYKTFGPFLMSFPFKAIHDLMPFSGPQILKLLIKFVNDTKAPD 360
DB 301 IVNSPQKEMNPSLFKVLYKTFGPFLMSFPFKAIHDLMPFSGPQILKLLIKFVNDTKAPD 360
QY 361 WQGYFYTVLLFVTAQCLQTLVLAHQYFHICFVSGMRKIKTAVIGAVYRKALVTNSARKSTV 420
DB 361 WQGYFYTVLLFVTAQCLQTLVLAHQYFHICFVSGMRKIKTAVIGAVYRKALVTNSARKSTV 420
QY 421 GEIYNLMSVDAQRMDLATYINMIMSAPLQVITLALYLMLNGPSVLGAVAWMLAMPVN 480
DB 421 GEIYNLMSVDAQRMDLATYINMIMSAPLQVITLALYLMLNGPSVLGAVAWMLAMPVN 480
QY 481 AVNMAKTKTYOVAMKSKDNRIKLMNEILNGIKVLKLYAMELAFKDYALIROBELKVLK 540
DB 481 AVNMAKTKTYOVAMKSKDNRIKLMNEILNGIKVLKLYAMELAFKDYALIROBELKVLK 540
QY 541 KSAVLSAVGTFTWCTPFLVALCTPAVYVITDENNNIIDAQTAFAVSLAFNLRPLNLP 600
DB 541 KSAVLSAVGTFTWCTPFLVALCTPAVYVITDENNNIIDAQTAFAVSLAFNLRPLNLP 600
QY 601 MVISSIVQASVSLRLRIPLSHELEPDSIERRPVKGGGNSITVRNATFTWARSDEPT 660
DB 601 MVISSIVQASVSLRLRIPLSHELEPDSIERRPVKGGGNSITVRNATFTWARSDEPT 660
QY 661 LMGITFSEIPEGALVAVVQVCGKSSLSALLAEMDKVEGHVAIKGSVAVYPQQAQITOND 720
DB 661 LMGITFSEIPEGALVAVVQVCGKSSLSALLAEMDKVEGHVAIKGSVAVYPQQAQITOND 720
QY 721 SLRENIILPGCOLLEPYRVSIVQACALLPDLIELPSGDRTEIGEKVNLSSGQKORVSLAR 780
DB 721 SLRENIILPGCOLLEPYRVSIVQACALLPDLIELPSGDRTEIGEKVNLSSGQKORVSLAR 780
QY 781 AVVSNADITYLPDDLSAVDAHVGHIPENNVIGPKMKLNKRIILVTHSMNLTPOVDYIV 840
DB 781 AVVSNADITYLPDDLSAVDAHVGHIPENNVIGPKMKLNKRIILVTHSMNLTPOVDYIV 840
QY 841 MSGGKISEMGSYQELIARDGAFAFELKRYASTEOQDAENGVYGVSPGKEAKOMENGM 900
DB 841 MSGGKISEMGSYQELIARDGAFAFELKRYASTEOQDAENGVYGVSPGKEAKOMENGM 900
QY 901 LVTDASAGKOLOROLSSSSSSYSGDISRHHNSTAELOKAEAKKEETWKLMEADKAOGTQYKL 960
DB 901 LVTDASAGKOLOROLSSSSSSYSGDISRHHNSTAELOKAEAKKEETWKLMEADKAOGTQYKL 960
QY 961 SVYWDYKAIAGLFTSPLSTFLMCHNVASALASNWLSLMTDDPIVNGQETHTKYRLSYG 1020
DB 961 SVYWDYKAIAGLFTSPLSTFLMCHNVASALASNWLSLMTDDPIVNGQETHTKYRLSYG 1020
```

```
QY 1021 ALGISQIAVFGISMAVSTIGIILASRCLHYVDLHSLIRSMSTFERTPSGNLYNRSEKL 1080
DB 1021 ALGISQIAVFGISMAVSTIGIILASRCLHYVDLHSLIRSMSTFERTPSGNLYNRSEKL 1080
QY 1081 DTVDMSIPEYIKAMMSLEFNVICACIVILATPIAIIIPPLGIYFVORFYVASSROL 1140
DB 1081 DTVDMSIPEYIKAMMSLEFNVICACIVILATPIAIIIPPLGIYFVORFYVASSROL 1140
QY 1141 KRLESVRSRPVYSHFNETLLGVSIVIRAFEEQERFIHOSDLKVDENOKAYYPSIVANMWA 1200
DB 1141 KRLESVRSRPVYSHFNETLLGVSIVIRAFEEQERFIHOSDLKVDENOKAYYPSIVANMWA 1200
QY 1201 VRLCEVNCIYLPALPAVISRHSLSAGLVGSVSYSLQVTTYLNLVRMSSEMETNIVA 1260
DB 1201 VRLCEVNCIYLPALPAVISRHSLSAGLVGSVSYSLQVTTYLNLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKAPMOIQTAPSSWPQVGRVFRNYCLRFBDDLPVLRHINVTINGG 1320
DB 1261 VERLKEYSETEKAPMOIQTAPSSWPQVGRVFRNYCLRFBDDLPVLRHINVTINGG 1320
QY 1321 EKVGIVRTGAGKSLTLGLFRINESAGEBIIIDGINIAKIGHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIVRTGAGKSLTLGLFRINESAGEBIIIDGINIAKIGHDLRFKTIIPQDPVLF 1380
QY 1381 SGLRMLNLDPPSQYSDSEVWTSLELAHLKDFVSALPDKLDHECAEGGENLSVGORQLVCL 1440
DB 1381 SGLRMLNLDPPSQYSDSEVWTSLELAHLKDFVSALPDKLDHECAEGGENLSVGORQLVCL 1440
QY 1441 ARALIRKTKILVDEAFAAADLETDDLIOSTIRTOPEDCVTLTAHRLNTIMYTRIVL 1500
DB 1441 ARALIRKTKILVDEAFAAADLETDDLIOSTIRTOPEDCVTLTAHRLNTIMYTRIVL 1500
QY 1501 DKGEIOEYGAPSDILOQRGLFYGMADKAGLV 1531
DB 1501 DKGEIOEYGAPSDILOQRGLFYGMADKAGLV 1531

RESULT 10
ABG61810
ID ABG61810 standard; protein; 1531 AA.
XX
AC ABG61810;
XX
DT 15-AUG-2002 (first entry)
XX
DE Prostate cancer-associated protein #11.
XX
KW Prostate cancer; prostate tumour tissue; human; mammal; cytosstatic.
XX
OS Mammalia.
XX
PN W0200230268-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US032045.
XX
PR 13-OCT-2000; 2000US-00687576.
PR 08-DEC-2000; 2000US-00733288.
PR 08-DEC-2000; 2000US-00733742.
PR 24-JAN-2001; 2001US-0263957P.
PR 16-MAR-2001; 2001US-026791P.
PR 16-MAR-2001; 2001US-0276888P.
PR 06-APR-2001; 2001US-0281922P.
PR 24-APR-2001; 2001US-0286214P.
PR 30-APR-2001; 2001US-00847046.
PR 04-MAY-2001; 2001US-0288589P.
XX
PA (BOSB-) BOS BIOTECHNOLOGY INC.
XX
PI Glah KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX
WP; 2002-471335/50.
```

DR N-PSDB; ABK92125.
XX Detecting a prostate cancer-associated transcript in a cell in a patient.
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.
XX
XX Claim 27; Page 309; 436pp; English.
XX
XX The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with prostate cancer-
CC associated polynucleotides (designated PC genes) that selectively
CC hybridize to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins
XX
XX Sequence 1531 AA;
SQ
Query Match 78.5%; Score 7860; DB 5; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MALRGFCSADGSPDLMNNTNTSNDPTKCPONTLVWPCPYLMACFPFYLYSRH 60
DB 1 MALRGFCSADGSPDLMNNTNTSNDPTKCPONTLVWPCPYLMACFPFYLYSRH 60
QY 61 DRGVIQWTPANKTKTALGFLMIYCMADLFYSFEMERSGIFLAPVFLVSTLGIITLLA 120
DB 61 DRGVIQWTPANKTKTALGFLMIYCMADLFYSFEMERSGIFLAPVFLVSTLGIITLLA 120
QY 121 TFLIQLERRKGVSSGIMLTFWLVALVCALAILRSKIMTALKEDAOVDLFRDITFYVYS 180
DB 121 TFLIQLERRKGVSSGIMLTFWLVALVCALAILRSKIMTALKEDAOVDLFRDITFYVYS 180
QY 181 LLLIQLVLSGFSRSPLEFSTIHDNPNCPRESSASFLSRITFWMTTGLIYVGYRPLEGSD 240
DB 181 LLLIQLVLSGFSRSPLEFSTIHDNPNCPRESSASFLSRITFWMTTGLIYVGYRPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVVKMKKCECAKTRKQPVVYSSKDPQOPKSSKYDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVVKMKKCECAKTRKQPVVYSSKDPQOPKSSKYDANEVEAL 300
QY 301 IVKSPQKEMNPSLFKVLTKTFGPFYFLMSFFPKAIDHLMFSGPQILKLLIKFVNDTKAPD 360
DB 301 IVKSPQKEMNPSLFKVLTKTFGPFYFLMSFFPKAIDHLMFSGPQILKLLIKFVNDTKAPD 360
QY 361 WQGYFTVLLFVTAQOTLVLAHQYFHCYSGMKRIKTAIVGAYYRKALVITNSARKSSTY 420
DB 361 WQGYFTVLLFVTAQOTLVLAHQYFHCYSGMKRIKTAIVGAYYRKALVITNSARKSSTY 420
QY 421 GEIVNLMSVDAQRFMDLATYINMIWAPLQVILALYLLMNLGSPVLAGAVWVLMVLPVN 480
DB 421 GEIVNLMSVDAQRFMDLATYINMIWAPLQVILALYLLMNLGSPVLAGAVWVLMVLPVN 480
QY 481 AVAMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAMELAFKQVLAIRQBELKVLK 540
DB 481 AVAMAKTKTYQVAHMKSKDNRIKLMNEILNGIKVLKYAMELAFKQVLAIRQBELKVLK 540
QY 541 KSAVYLSAVGTFVWCPFLVALCTPAYVTIIDENNIIIDACTAVSIALFPPLNIIP 600
DB 541 KSAVYLSAVGTFVWCPFLVALCTPAYVTIIDENNIIIDACTAVSIALFPPLNIIP 600
QY 601 MVVSSIVQASVSLKRLRIFLSHELEBPSIERRPVKGGGTNSITVANAFTTARSDPT 660
DB 601 MVVSSIVQASVSLKRLRIFLSHELEBPSIERRPVKGGGTNSITVANAFTTARSDPT 660

DB 601 MVVSSIVQASVSLKRLRIFLSHELEBPSIERRPVKGGGTNSITVANAFTTARSDPT 660
QY 661 LNTGTFSPGALVAVVGVGCGKSSLSLALLAEMDVBEHVAIKGSVAVYPOQAWIQND 720
DB 661 LNTGTFSPGALVAVVGVGCGKSSLSLALLAEMDVBEHVAIKGSVAVYPOQAWIQND 720
QY 721 SLRENILFGQLEBPYRSYIOACALLPDEIILPSGRTETIGEKVNLSSGQKORVSLAR 780
DB 721 SLRENILFGQLEBPYRSYIOACALLPDEIILPSGRTETIGEKVNLSSGQKORVSLAR 780
QY 781 AVYSNADIVLPDPLSADVAHVGKHIPENYIGPKMKLKNKTRILVTHSMSYLPQVDVIY 840
DB 781 AVYSNADIVLPDPLSADVAHVGKHIPENYIGPKMKLKNKTRILVTHSMSYLPQVDVIY 840
QY 841 MSGGKISEMSYQELARDQAPAPLFTYASTEOBDAENGTGVSQPKKAKOMENGM 900
DB 841 MSGGKISEMSYQELARDQAPAPLFTYASTEOBDAENGTGVSQPKKAKOMENGM 900
QY 901 LVYDSAGKQORQOLSSSSYSGDISSRHNSYAELOKAEKKEETKLMKADKQOTGVKL 960
DB 901 LVYDSAGKQORQOLSSSSYSGDISSRHNSYAELOKAEKKEETKLMKADKQOTGVKL 960
QY 961 SVYWDYMKALGLFISFLSIFLPMCNHVSALASNYWMLSLMTDDPIVNGTOHTKVRLSVYG 1020
DB 961 SVYWDYMKALGLFISFLSIFLPMCNHVSALASNYWMLSLMTDDPIVNGTOHTKVRLSVYG 1020
QY 1021 ALGISOGIAVGYSMASVIGIILASRCLHVDLHSLIRSPMSFPERTSGNLVNRFSKEL 1080
DB 1021 ALGISOGIAVGYSMASVIGIILASRCLHVDLHSLIRSPMSFPERTSGNLVNRFSKEL 1080
QY 1081 DTVDMSIPEVINKFMGSLFNVIGACTVILATPIAIIIPDLGIYFPVQRFVASSROL 1140
DB 1081 DTVDMSIPEVINKFMGSLFNVIGACTVILATPIAIIIPDLGIYFPVQRFVASSROL 1140
QY 1141 KRLESVRSRPVSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLESVRSRPVSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLCVCNCIVLFAALPAVISRHSLSAGLVGSVSYSLQVTTYLANMLVRMSSEMETNIVA 1260
DB 1201 VRLCVCNCIVLFAALPAVISRHSLSAGLVGSVSYSLQVTTYLANMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSFETKEKAPMOIQETAPSSWPQGRVFEFNRYCLRYREDLDPEVLRHINVTINGG 1320
DB 1261 VERLKEYSFETKEKAPMOIQETAPSSWPQGRVFEFNRYCLRYREDLDPEVLRHINVTINGG 1320
QY 1321 EKVGIYGRGAGKSSLTGLFRINESABGEIITDGINIAKIGHDLRFKTIIPQDPVLF 1380
DB 1321 EKVGIYGRGAGKSSLTGLFRINESABGEIITDGINIAKIGHDLRFKTIIPQDPVLF 1380
QY 1381 SSSLRNMLDPFSQYSDPEVWTSLELAHLKDPVSALPDKLDHECABEGENLSVGOROLVCL 1440
DB 1381 SSSLRNMLDPFSQYSDPEVWTSLELAHLKDPVSALPDKLDHECABEGENLSVGOROLVCL 1440
QY 1441 ARLALKTKTILVDEATAVADLETDDLIOSTTRTOEDCTVLTARLNTIMTYRVIYL 1500
DB 1441 ARLALKTKTILVDEATAVADLETDDLIOSTTRTOEDCTVLTARLNTIMTYRVIYL 1500
QY 1501 DKGEIOEYGAPSDLQORGFLYSMAKDAGLV 1531
DB 1501 DKGEIOEYGAPSDLQORGFLYSMAKDAGLV 1531
RESULT 11
ID ABM35012 standard; protein; 1531 AA.
XX ABM35012:
AC ABM35012:
XX
DT 08-OCT-2003 (first entry)
XX
XX Cancer based on CYP3A5 related protein seq ID NO:678.
XX


```
|||||
Db      1441 ABALRLKTKILVDEAPAAVDLETFDDLIQSTIRTFQFECYTLTAHLNLTIMDTTRYIVL 1500
Qy      1501 DKGEIQEYGA PSDLLQQRGCLFYSMAXDAGLV 1531
      |||||
Db      1501 DKGEIQEYGA PSDLLQQRGCLFYSMAXDAGLV 1531

RESULT 12
ADB20865
ID      ADB20865 standard; protein; 1531 AA.
XX
XX      ADB20865;
XX
XX      20-NOV-2003 (first entry)
XX
XX      MRP1 based cancer related protein SEQ ID NO:678.
XX
XX      irinotecan; colorectal cancer; cervical cancer; gastric cancer;
KM      lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
KM      variant allele; multidrug resistance protein 1; MRP1; cytosolic.
XX
XX      Unidentified.
OS
XX      MO2003013533-A2.
XX
XX      20-FEB-2003.
XX
XX      23-JUL-2002; 2002MO-EP008200.
XX
XX      23-JUL-2001; 2001EP-00117608.
PR      24-MAY-2002; 2002EP-00011710.
XX
XX      (EPID-) EPIDAUROS BIOECHANOLOGIE AG.
XX
PI      Heinrich G, Kerb R;
XX
XX      WPI; 2003-354397/33.
XX
XX      Use of irinotecan or its derivative for preparation of a pharmaceutical
PT      composition for treating cancer in a subject having a genome with a
PT      variant allele comprising a multidrug resistance protein 1
PT      polynucleotide.
XX
XX      Disclosure, SEQ ID NO 678; 100pp; English.
XX
XX      The present invention describes a method for the use of irinotecan (I) or
CC      its derivative for the preparation of a pharmaceutical composition for
CC      treating colorectal, cervical, gastric, lung, ovarian or pancreatic
CC      cancer, or malignant glioma in a subject having a genome with a variant
CC      allele which comprises a multidrug resistance protein 1 (MRP1)
CC      polynucleotide (II). (I) has cytostatic activity. (I) or its derivative
CC      can be used for the preparation of a pharmaceutical composition for
CC      treating colorectal, cervical, gastric, lung, ovarian or pancreatic
CC      cancer, or malignant glioma in a subject, where the subject is a human
CC      (preferably African or Asian) or a mouse. The present sequence represents
CC      a sequence which is used in the exemplification of the present invention.
XX
XX      Sequence 1531 AA;

Query Match      78.5%; Score 7860; DB 6; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MALRGFCSADGSDPLMDMNTVNTSNPDFTKCFQNTLVVWPCFYLMACPFYFLYSRH 60
Db      1 MALRGFCSADGSDPLMDMNTVNTSNPDFTKCFQNTLVVWPCFYLMACPFYFLYSRH 60
Qy      61 DRGYIQMTPLNKTALGFLMTVCMADLYSTFERSRGIFLAPVFLVSTLLGITTLA 120
Db      61 DRGYIQMTPLNKTALGFLMTVCMADLYSTFERSRGIFLAPVFLVSTLLGITTLA 120
Qy      121 TFLIQERRRGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDFRDIITFYVFS 180
Db      121 TFLIQERRRGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDFRDIITFYVFS 180

|||||
Db      121 TFLIQERRRGVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAQVDFRDIITFYVFS 180
Qy      181 LLLIQVLVSCFSDRSPLFSETHIDNFCPESSASFLSRITFWMTGLIVRGYRPLEGSD 240
      |||||
Db      181 LLLIQVLVSCFSDRSPLFSETHIDNFCPESSASFLSRITFWMTGLIVRGYRPLEGSD 240
Qy      241 LMSLNKEDTSEOVVPUYKRWKKECAKTRQPKVYVSSKDPQPKSSVVDANEVEAL 300
      |||||
Db      241 LMSLNKEDTSEOVVPUYKRWKKECAKTRQPKVYVSSKDPQPKSSVVDANEVEAL 300
Qy      301 IVKSPQKEMNPSLFKVLYKTFGPFYLMSPFFKAIHDLMSGPOILKLIKFNVDTRAPD 360
      |||||
Db      301 IVKSPQKEMNPSLFKVLYKTFGPFYLMSPFFKAIHDLMSGPOILKLIKFNVDTRAPD 360
Qy      361 WQGYFTVLLFVTAQLOTLVLAHQYFHCYVSGRIKTAIVGAAYRKALVITNSARKSTV 420
      |||||
Db      361 WQGYFTVLLFVTAQLOTLVLAHQYFHCYVSGRIKTAIVGAAYRKALVITNSARKSTV 420
Qy      421 GEIVNLSVDAQRFMDLATYINMWSAPLOVILALYLMNLGPSVLGVAVMVLAHPVN 480
      |||||
Db      421 GEIVNLSVDAQRFMDLATYINMWSAPLOVILALYLMNLGPSVLGVAVMVLAHPVN 480
Qy      481 AVNANKTKTYQVAHMSKDNRIKLMNEILNGIKVLYKLYAMELAFKDKVLAIROBELKVLK 540
      |||||
Db      481 AVNANKTKTYQVAHMSKDNRIKLMNEILNGIKVLYKLYAMELAFKDKVLAIROBELKVLK 540
Qy      541 KSAVLSAVGTFTVTCPTPLVALCTPAVYVITIDENNIIDAQTAFAVSLAFNILRPLNLP 600
      |||||
Db      541 KSAVLSAVGTFTVTCPTPLVALCTPAVYVITIDENNIIDAQTAFAVSLAFNILRPLNLP 600
Qy      541 KSAVLSAVGTFTVTCPTPLVALCTPAVYVITIDENNIIDAQTAFAVSLAFNILRPLNLP 600
Qy      601 MVISSIVQASVSLKRLRIFLSHEELPDSIERRPVKGCGGTSITVNAATFTMARSPPT 660
      |||||
Db      601 MVISSIVQASVSLKRLRIFLSHEELPDSIERRPVKGCGGTSITVNAATFTMARSPPT 660
Qy      661 LINGITTSIPREGALVAVVGVQVCGKSSLSLALBMDKVEGHVAIKSVAVPOQAWQND 720
      |||||
Db      661 LINGITTSIPREGALVAVVGVQVCGKSSLSLALBMDKVEGHVAIKSVAVPOQAWQND 720
Qy      721 SLRENILFGCQLEBPYRYSVYQACALLPDLEILPSGDRTEIGRGNVLSGGQKQVSLAR 780
      |||||
Db      721 SLRENILFGCQLEBPYRYSVYQACALLPDLEILPSGDRTEIGRGNVLSGGQKQVSLAR 780
Qy      781 AVYSNADITYLPDPLSADVAHVGKHIPENVYIGPKMKKNTKRLLVTHSMGYLPQVDYIIV 840
      |||||
Db      781 AVYSNADITYLPDPLSADVAHVGKHIPENVYIGPKMKKNTKRLLVTHSMGYLPQVDYIIV 840
Qy      841 MSGGKISEMGSYQELLARDGAFAPFLRTYASTROBDAENGVYVSGPKEKAKOMENGM 900
      |||||
Db      841 MSGGKISEMGSYQELLARDGAFAPFLRTYASTROBDAENGVYVSGPKEKAKOMENGM 900
Qy      901 LVYDSAGKQLOROLQSSSSSYSGDISRHHNSTALQRAEAKERTWKLMEADKAQOTGVKL 960
      |||||
Db      901 LVYDSAGKQLOROLQSSSSSYSGDISRHHNSTALQRAEAKERTWKLMEADKAQOTGVKL 960
Qy      961 SVYWDYKAIKGLFISFLSIFLFCNHNVSALASNYWLSLWTDPIVNGTOGHTYKRLSVYG 1020
      |||||
Db      961 SVYWDYKAIKGLFISFLSIFLFCNHNVSALASNYWLSLWTDPIVNGTOGHTYKRLSVYG 1020
Qy      1021 ARGISQIIVFQVSMVSTIGLILASRCLAVDLDLHSLRSMSPFERTPSGNLVNRRSKEL 1080
      |||||
Db      1021 ARGISQIIVFQVSMVSTIGLILASRCLAVDLDLHSLRSMSPFERTPSGNLVNRRSKEL 1080
Qy      1081 DTVDSMIPEVIKMFMSLFFNVIGACVILLATPAAIIIPPLDLIFVFVORFYVASSROL 1140
      |||||
Db      1081 DTVDSMIPEVIKMFMSLFFNVIGACVILLATPAAIIIPPLDLIFVFVORFYVASSROL 1140
Qy      1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEOERFIHQSDLKVDENKAYYPSIVANRWLA 1200
      |||||
Db      1141 KRLSVSRSPVYSHFNETLLGVSVIRAFEOERFIHQSDLKVDENKAYYPSIVANRWLA 1200
Qy      1201 VRLCEGVNCIVFAALFAVYSRHSLSAGLVGVSVSLSQVTTYLVNMLVNMSSMEETNIYA 1260
      |||||
Db      1201 VRLCEGVNCIVFAALFAVYSRHSLSAGLVGVSVSLSQVTTYLVNMLVNMSSMEETNIYA 1260
```

Db	1201	VRLEBCVNCIYLPALPAFVIVSRHSISAGLVGLSVSYSLQVTTYYLNMVLRMSSEMETNIVA	1266
Qy	1361	VERLKEVSETEKEAPWQIOETAPPSWBPQVGRVEFRNVCYAFYREDDJDFVLRHINVTNGG	1320
Db	1261	VERLKEVSETEKEAPWQIOETAPPSWBPQVGRVEFRNVCYAFYREDDJDFVLRHINVTNGG	1320
Qy	1321	EKVGIVGTGAGKSLTYGLFRINSAGCEIITDGINAKIGLHDLRFKIIITIPDDPLF	1360
Db	1321	EKVGIVGTGAGKSLTYGLFRINSAGCEIITDGINAKIGLHDLRFKIIITIPDDPLF	1360
Qy	1381	SGSLRPNIDPPSQYSDDEEVTWSLEIAHLKDPVSALPDXLHDCACGGENISVGORQIVCL	1440
Db	1381	SGSLRPNIDPPSQYSDDEEVTWSLEIAHLKDPVSALPDXLHDCACGGENISVGORQIVCL	1440
Qy	1441	ARALRRTKIIIVLEATAAVDLETDDLIQSTIRTFEDCYVLTIAHRLANTIMDTRYVVL	1500
Db	1441	ARALRRTKIIIVLEATAAVDLETDDLIQSTIRTFEDCYVLTIAHRLANTIMDTRYVVL	1500
Qy	1501	DKGEIOEVGAPSDLIQQRGLFYSMAKDGLV	1531
Db	1501	DKGEIOEVGAPSDLIQQRGLFYSMAKDGLV	1531

RESULT 13
 ADB87954
 ID ADB87954 standard; protein; 1531 AA.
 XX
 AC ADB87954;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Human UGT1A1 protein sequence SEQ ID NO:678.
 XX
 KW irinotecan; cancer; UGT1A1; cytosolic; topoisomerase I inhibitor;
 KW colorectal cancer; cervical cancer; gastric cancer; lung cancer;
 KW ovarian cancer; pancreatic cancer; malignant glioma;
 KW uridine diphosphate glycosyltransferase1 member A1.
 XX
 OS Homo sapiens.
 OS
 PN WO2003013536-A2.
 PN
 PD 20-FEB-2003.
 PD
 PF 23-JUL-2002; 2002MO-EP008217.
 PF
 PR 23-JUL-2001; 2001EP-00117608.
 PR
 PR 24-MAY-2002; 2002EP-00011710.
 XX
 PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
 PI
 PI Heinrich G, Kerb R;
 PI
 DR WPI; 2003-289896/28.
 DR
 PT Use of irinotecan to treat cancer patient by determining if patient has
 PT variant alleles of UGT1A1 gene, administering increased/decreased amounts
 PT of irinotecan based on increased/decreased levels of UGT1A1 gene product.
 PT
 PS Disclosure; SEQ ID NO 678; 107bp; English.
 PS
 CC The invention relates to the novel use of irinotecan to treat a patient
 CC suffering from cancer. This involves determining if the patient has one
 CC or more variant alleles of the UGT1A1 gene, and if the patient has one
 CC or more of such variant alleles, irinotecan is administered in an increased
 CC or decreased amount in comparison to the amount that is administered
 CC without regard to the patient's alleles in the UGT1A1 gene. The invention
 CC has cytosolic activity. A composition of the invention acts as a
 CC topoisomerase I inhibitor. The method is useful for treating a patient,
 CC an animal e.g. mouse or a human, preferably African or Asian, suffering
 CC from cancer such as colorectal, cervical, gastric cancer, lung, ovarian,
 CC pancreatic cancer or malignant glioma. The present sequence is udes in
 CC the exemplification of the invention.

XX	Sequence	1531 AA,	78.5%; Score 7860; DB 7; Length 1531;
Query Match	78.5%; Score 7860; DB 7; Length 1531;		
Best Local Similarity	100.0%; Prod. No. 0;		
Matches 1531;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	MAIRGFCGADSDPLMDMNVMTMNTSNDPFTKCFONTLWVWPCEYLMACPEPFYLSRH	60
DB	1	MAIRGFCGADSDPLMDMNVMTMNTSNDPFTKCFONTLWVWPCEYLMACPEPFYLSRH	60
QY	61	DRGYIQMTPLAKTKTALGFLLMIYCMADLFFSFMEBSRGIFLAVFLVSPILGITTLLA	120
DB	61	DRGYIQMTPLAKTKTALGFLLMIYCMADLFFSFMEBSRGIFLAVFLVSPILGITTLLA	120
QY	121	TFLLQLEERKVOSSGIMLTFMLVALCALALTSKIMTALKEAOVDLPDIFYYVFFS	180
DB	121	TFLLQLEERKVOSSGIMLTFMLVALCALALTSKIMTALKEAOVDLPDIFYYVFFS	180
QY	181	LLLIQLVLSCFSDSPLESETIHDNPNCPESSASFLSRTFFWITGILVRCYROPLEBSD	240
DB	181	LLLIQLVLSCFSDSPLESETIHDNPNCPESSASFLSRTFFWITGILVRCYROPLEBSD	240
QY	241	LMSLNKEDTSQQVPLVYKMKKECATRKQPVKVVYSSKQDPAQKSSKTDANEYBAL	300
DB	241	LMSLNKEDTSQQVPLVYKMKKECATRKQPVKVVYSSKQDPAQKSSKTDANEYBAL	300
QY	301	IVKSPQKMNNSLFRVLYKTEGPFELMSFFPKALHDLMFSGPQILKLLIKFVNDTRAPD	360
DB	301	IVKSPQKMNNSLFRVLYKTEGPFELMSFFPKALHDLMFSGPQILKLLIKFVNDTRAPD	360
QY	361	WQGYFYTLFVTACTQLTLVHQFPHLCFVSGMKIKTAVIGAVVRKALVITNSARKSSTV	420
DB	361	WQGYFYTLFVTACTQLTLVHQFPHLCFVSGMKIKTAVIGAVVRKALVITNSARKSSTV	420
QY	421	GEIVNLSVDAQRFMDLATYINMTWSAPLOYIALYILMTNLGFSVLGAVAVMTLMPVN	480
DB	421	GEIVNLSVDAQRFMDLATYINMTWSAPLOYIALYILMTNLGFSVLGAVAVMTLMPVN	480
QY	481	AVMAKTKTYOVAMHMSKDNRIKLMNELLNIGIKYLKLYAMELAKKDYALROBELVYTK	540
DB	481	AVMAKTKTYOVAMHMSKDNRIKLMNELLNIGIKYLKLYAMELAKKDYALROBELVYTK	540
QY	541	KSAYLSAVGTETWCTPEPLVALCTFAVYVITDENNNILDAQFAFYSLALFNILRPEPLILP	600
DB	541	KSAYLSAVGTETWCTPEPLVALCTFAVYVITDENNNILDAQFAFYSLALFNILRPEPLILP	600
QY	601	MVSSIVQASVSLKRLRIFLSHEELPDSIERRPVKQGGGNTSITVENATFTMARSDPT	660
DB	601	MVSSIVQASVSLKRLRIFLSHEELPDSIERRPVKQGGGNTSITVENATFTMARSDPT	660
QY	661	INGITFSTPEBALVAVVQVQCCGSSLSLALAEMLDVEBGVVALKGSVAAYVPOQAMQND	720
DB	661	INGITFSTPEBALVAVVQVQCCGSSLSLALAEMLDVEBGVVALKGSVAAYVPOQAMQND	720
QY	721	SLRENIIEGCOLEBPYRSVIOACALPDEILPSGRTIEGEGVNLSSGQXQVSLAR	780
DB	721	SLRENIIEGCOLEBPYRSVIOACALPDEILPSGRTIEGEGVNLSSGQXQVSLAR	780
QY	781	AVYSNADIYLPDDPLSAVDHVGKHIPEENVIGPKGMLKNKTRILVTHSMSYLPOVDYIIV	840
DB	781	AVYSNADIYLPDDPLSAVDHVGKHIPEENVIGPKGMLKNKTRILVTHSMSYLPOVDYIIV	840
QY	841	MSGKISMSYSQELIARDAFAFRLRTYASTBOBDAEENGVTGVSQPGKEAQMENGMM	900
DB	841	MSGKISMSYSQELIARDAFAFRLRTYASTBOBDAEENGVTGVSQPGKEAQMENGMM	900
QY	901	LYTBSAGQOLROLOSSSSSSSGDISRBHNSYAELOKAEKKEETWKLMEADKAOTGOYKL	960
DB	901	LYTBSAGQOLROLOSSSSSSSGDISRBHNSYAELOKAEKKEETWKLMEADKAOTGOYKL	960
QY	961	SVYMDYMKALGFLISFLIFLFCNHNYSALASNWLSLMTDPIVNGTOEHTKYRLSYGG	1020

Db 961 SVYWDYKAIAGLFIISFISLIFLMCNHVSALASNYWLSLMTDDPIVNGTQHTKVLRSVYG 1020

Qy 1021 ALGISGIAVFGVSMASIGIILASRCLADLHSLIRSPMSFEERPPSGNLVNRSEKEL 1080

Db 1021 ALGISGIAVFGVSMASIGIILASRCLADLHSLIRSPMSFEERPPSGNLVNRSEKEL 1080

Qy 1081 DTVDSMIPFVYIKMFMSLFFNVIGACIYLATPAAIIIPPLGLIYFVORFYVASSROL 1140

Db 1081 DTVDSMIPFVYIKMFMSLFFNVIGACIYLATPAAIIIPPLGLIYFVORFYVASSROL 1140

Qy 1141 KRLESVRSRSPVYSHFNFTLLGVSVIRAFEEQERFIHQSDLKVDENQAYYPSIVANRWLA 1200

Db 1141 KRLESVRSRSPVYSHFNFTLLGVSVIRAFEEQERFIHQSDLKVDENQAYYPSIVANRWLA 1200

Qy 1201 VRLCEGNCIVLPAALPAVVISRHSLSAGLVGLSVSYSLQVTTIYLMVLRMSSEMETIYA 1260

Db 1201 VRLCEGNCIVLPAALPAVVISRHSLSAGLVGLSVSYSLQVTTIYLMVLRMSSEMETIYA 1260

Qy 1261 VERLKESETEKAPMOIOETAPSSWPQVGFVFRNYCRLYREDDLPVLRHINVTINGG 1320

Db 1261 VERLKESETEKAPMOIOETAPSSWPQVGFVFRNYCRLYREDDLPVLRHINVTINGG 1320

Qy 1321 EKVGIVGRTGAGKSLTLGLFRINESABEIIIDGINIAKIGLHDLRFKITTIIPODPVLF 1380

Db 1321 EKVGIVGRTGAGKSLTLGLFRINESABEIIIDGINIAKIGLHDLRFKITTIIPODPVLF 1380

Qy 1381 SGSLRANLDPFSQYSDEEWTSLFLAHLKDFVSALPKLDHECAGEGENSEVQOROLVCL 1440

Db 1381 SGSLRANLDPFSQYSDEEWTSLFLAHLKDFVSALPKLDHECAGEGENSEVQOROLVCL 1440

Qy 1441 ARRLKRTKILVDEATPAVDLETDDLIOSTIRTOPEDCVTLTAHRLNTIMYTRYIVL 1500

Db 1441 ARRLKRTKILVDEATPAVDLETDDLIOSTIRTOPEDCVTLTAHRLNTIMYTRYIVL 1500

Qy 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531

Db 1501 DKGEIOEYGAPSDLLQORGLFYSMAKDAGLV 1531

RESULT 14

ID ADB96937 standard; protein; 1531 AA.

XX ADB96937;

AC 04-DEC-2003 (first entry)

DT Human MDR1 related protein sequence SEQ ID NO:678.

DE

XX

KM irinotecan; colorectal cancer; cervical cancer; gastric cancer;

KM lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;

KM multidrug resistance 1; MDR1; cytostatic; human; CYP3A5; MRP1; MDR1;

KM TOP1.

XX

OS Homo sapiens.

XX

PN WO2003013537-A2.

XX

PD 20-FEB-2003.

PF 23-JUL-2002; 2002WO-EP008218.

XX

PR 23-JUL-2001; 2001EP-00117608.

PR 24-MAY-2002; 2002EP-00011710.

XX

PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.

XX

PI Heinrich G, Kerb R;

XX

DR WPI; 2003-268145/26.

PT New use of irinotecan for preparation of pharmaceutical compositions for treating cancer in subject having genome with variant allele comprising

PT multidrug resistance 1 polynucleotide.

XX

XX Disclosure; SEQ ID NO 678; 130bp; English.

XX

CC The invention relates to the novel use of irinotecan or its derivative

CC for the preparation of pharmaceutical compositions for treating

CC colorectal, cervical, gastric, lung, ovarian or pancreatic cancer, or

CC malignant glioma in a subject having a genome with a variant allele which

CC comprises a multidrug resistance 1 (MDR1) polynucleotide. A composition

CC of the invention has cytostatic activity. The invention is useful for the

CC preparation of pharmaceutical compositions for treating colorectal,

CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant

CC glioma in a subject (preferably human, more preferably African or Asian)

CC or a mouse. The present sequence is used in the exemplification of the

CC invention.

XX

XX Sequence 1531 AA;

XX

SQ

Query Match 78.5%; Score 7860; DB 7; Length 1531;

Best Local Similarity 100.0%; Pred. No.0;

Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MALRGFCSDAGSDPLMDMNTWNTSNPDFTKCFQNTVLVWPCFYLMACFPFYLYLSRH 60

Db 1 MALRGFCSDAGSDPLMDMNTWNTSNPDFTKCFQNTVLVWPCFYLMACFPFYLYLSRH 60

Qy 61 DRGYIOMTPINKTKYALGFLMIVCWADLFYSFWEBSRGIFLAPVFLVSTLIGITTLA 120

Db 61 DRGYIOMTPINKTKYALGFLMIVCWADLFYSFWEBSRGIFLAPVFLVSTLIGITTLA 120

Qy 121 TFLIOERRKGVSSGIMTLFWLVALVCAALIRSKIMTALKEADAQVDRDITFVYVS 180

Db 121 TFLIOERRKGVSSGIMTLFWLVALVCAALIRSKIMTALKEADAQVDRDITFVYVS 180

Qy 181 LLLIQVLSCFSDRSLPSETTIDHPNCPRESSAFISRTIFMMITGLIYVGRQPLEGSD 240

Db 181 LLLIQVLSCFSDRSLPSETTIDHPNCPRESSAFISRTIFMMITGLIYVGRQPLEGSD 240

Qy 241 LMSLNKEDTSEQVPLVKNWKKCAKTRKQPVKVYSSKDPAPKRESSKVDANEVEAL 300

Db 241 LMSLNKEDTSEQVPLVKNWKKCAKTRKQPVKVYSSKDPAPKRESSKVDANEVEAL 300

Qy 301 IVKSPQKWNPSLFXLYKTFPGYFLMSFFPKAIDHLMFSGQIILKLIKPNDRKAPD 360

Db 301 IVKSPQKWNPSLFXLYKTFPGYFLMSFFPKAIDHLMFSGQIILKLIKPNDRKAPD 360

Qy 361 WQGYFTVLLFVYACQOTVLHQYFHCFSGMRKTAIVGAYYRKALVITNSAKSSTV 420

Db 361 WQGYFTVLLFVYACQOTVLHQYFHCFSGMRKTAIVGAYYRKALVITNSAKSSTV 420

Qy 421 GEIVNLSYDAQGFMDLATYINNIWSAPLOVILALYLMNLGSPVLAQVAVVLMVNV 480

Db 421 GEIVNLSYDAQGFMDLATYINNIWSAPLOVILALYLMNLGSPVLAQVAVVLMVNV 480

Qy 481 AVNAAKTKTYOYAHMNSKNNRIKLMNLIINGIKVLKYAMELAFKQKVLAIROBELKVLK 540

Db 481 AVNAAKTKTYOYAHMNSKNNRIKLMNLIINGIKVLKYAMELAFKQKVLAIROBELKVLK 540

Qy 541 KSAVLSAVGFTFWVCTPFLVALCTPAVYVYTIIDENILIDQOTAFVSLAFNIIAFPLNIIP 600

Db 541 KSAVLSAVGFTFWVCTPFLVALCTPAVYVYTIIDENILIDQOTAFVSLAFNIIAFPLNIIP 600

Qy 601 MYISSIVQASVSLKRLRIFLSHEELEPDSIERRPVQGGGJNSITVRNATFTWASDPT 660

Db 601 MYISSIVQASVSLKRLRIFLSHEELEPDSIERRPVQGGGJNSITVRNATFTWASDPT 660

Qy 661 LAGITFSIEGALVAVVGVGGKSLSALLAEMKVGSHVAKSVAVVPOQAIQND 720

Db 661 LAGITFSIEGALVAVVGVGGKSLSALLAEMKVGSHVAKSVAVVPOQAIQND 720

Qy 721 SLRENTLFGCOLPEEPYRSVIOACALLPDLLELPSSDRTEIGSKVNLGGQKQKRVSLAR 780

Db 721 SLRENTLFGCOLPEEPYRSVIOACALLPDLLELPSSDRTEIGSKVNLGGQKQKRVSLAR 780

QY 781 AVYSNMDIYLFDDPLSAVDHVGKHFENYIGPKGMLKNKTRILVTHSMSTLPQVDYIIV 840
DB 781 AVYSNMDIYLFDDPLSAVDHVGKHFENYIGPKGMLKNKTRILVTHSMSTLPQVDYIIV 840
QY 841 MSGGKISEMSYQELRLRDGAFAPFLTYASTEBODAEENGVTGVSQPKGAOMENGM 900
DB 841 MSGGKISEMSYQELRLRDGAFAPFLTYASTEBODAEENGVTGVSQPKGAOMENGM 900
QY 901 LMTDSAGKOLOROLSSSSSYSGDISRHNSPAELQXAEAKEEELWKMEADKATQGVKL 960
DB 901 LMTDSAGKOLOROLSSSSSYSGDISRHNSPAELQXAEAKEEELWKMEADKATQGVKL 960
QY 961 SVYWDYKAIQGLFISFISIFLMCNHVSALASNYMLSLMTDDPIVNGTOEHTKRLSVYG 1020
DB 961 SVYWDYKAIQGLFISFISIFLMCNHVSALASNYMLSLMTDDPIVNGTOEHTKRLSVYG 1020
QY 1021 ALGISGIAVFGSMANVSIGIILASRCLHVDLHSLRSPMSFEERTPSGNLVNRFSEKEL 1080
DB 1021 ALGISGIAVFGSMANVSIGIILASRCLHVDLHSLRSPMSFEERTPSGNLVNRFSEKEL 1080
QY 1081 DTVDMSIPEYIKMFGSLFNVTGACIVTLATPIAIIIPPLGIYFFVOQFYASSROL 1140
DB 1081 DTVDMSIPEYIKMFGSLFNVTGACIVTLATPIAIIIPPLGIYFFVOQFYASSROL 1140
QY 1141 KRLESVSRSPYSHFNETLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLESVSRSPYSHFNETLLGVSVIRAFEEQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLCEVNCIYLPALFAVISRHSLSAGLVSVYSIQVTTYINWLVNRSSEMETNIVA 1260
DB 1201 VRLCEVNCIYLPALFAVISRHSLSAGLVSVYSIQVTTYINWLVNRSSEMETNIVA 1260
QY 1261 VERLKEYSRTEKEAPMOIOETAPSPSPQVRVEPNYCLARYEDLPVLRIHIVTINGG 1320
DB 1261 VERLKEYSRTEKEAPMOIOETAPSPSPQVRVEPNYCLARYEDLPVLRIHIVTINGG 1320
QY 1321 EKVGIVRTGAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLREKTIIPQDPVLF 1380
DB 1321 EKVGIVRTGAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLREKTIIPQDPVLF 1380
QY 1381 SSGIRKMLDPPSQYSDEBWTSLBLAHKDFVSALPDKLDHECABGGENSVGQROLVCL 1440
DB 1381 SSGIRKMLDPPSQYSDEBWTSLBLAHKDFVSALPDKLDHECABGGENSVGQROLVCL 1440
QY 1441 ABALIRKTKILVLEATPAVDLETDLIQSTIRFOEDCTVLTIAHRLNTIMDTRYIVL 1500
DB 1441 ABALIRKTKILVLEATPAVDLETDLIQSTIRFOEDCTVLTIAHRLNTIMDTRYIVL 1500
QY 1501 DKGEIOEGAPSDLLQORGLFYSMADAGLV 1531
DB 1501 DKGEIOEGAPSDLLQORGLFYSMADAGLV 1531

RESULT 15
ADB92128
ID ADB92128 standard; protein, 1531 AA.

AC ADB92128;
DT 04-DEC-2003 (first entry)

DE Human MDR1 related protein sequence SEQ ID NO:678.

XX itinotecan; colorectal cancer; cervical cancer; gastric cancer;
XX lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
XX multidrug resistance 1; MDR1; cytosolic; human; UGT1A1; MRP1; Top1.

OS Homo sapiens.

PN W02003013535-A2.

XX 20-FEB-2003.

PD

XX 23-JUL-2002; 2002WO-EP008220.
PF 23-JUL-2001; 2001BP-00117608.
PR 24-MAY-2002; 2002BP-00011710.
PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
PI Heinrich G, Kerb R;
XX WPI; 2003-342400/32.
DR
XX
XX
PT New use of itinotecan for preparation of pharmaceutical compositions for
PT treating cancer in subject having genome with variant allele comprising
PT multidrug resistance 1 polynucleotide.
XX
XX
PS Disclosure; SEQ ID NO 678; 104pp; English.
XX
XX
CC The invention relates to a novel use of itinotecan or its derivative for
CC the preparation of a pharmaceutical composition for treating colorectal,
CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant
CC glioma in a subject having a genome with a variant allele which comprises
CC a multidrug resistance 1 (MDR1) polynucleotide. A composition of the
CC invention has cytostatic activity. The present sequence is used in the
CC exemplification of the invention.
XX
XX
SQ Sequence 1531 AA;

Query Match 78.5%; Score 7860; DB 7; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCSDAGSDPLMDMNTVNTSNDFPKCFQNTLVVWPCFYLMACFPFYFLYSRH 60
DB 1 MALRGFCSDAGSDPLMDMNTVNTSNDFPKCFQNTLVVWPCFYLMACFPFYFLYSRH 60
QY 61 DRGYIQTPLNKTKTALGFLIMIVCWADLFYSWERSRGIFLAPVFLVSTLLGITTLLA 120
DB 61 DRGYIQTPLNKTKTALGFLIMIVCWADLFYSWERSRGIFLAPVFLVSTLLGITTLLA 120
QY 121 TPLIQLERRKGVSSGIMLTFWLVALCALALIRSKIMTLAKEDAOVDLFRDITFYVFS 180
DB 121 TPLIQLERRKGVSSGIMLTFWLVALCALALIRSKIMTLAKEDAOVDLFRDITFYVFS 180
QY 181 LLLIQVLSCFSDRSPLFSETIHDNPPCPSSASFSLRITFMWITGLIVRGYQPLEGSD 240
DB 181 LLLIQVLSCFSDRSPLFSETIHDNPPCPSSASFSLRITFMWITGLIVRGYQPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVKYVSSKDPAPKSSKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVKYVSSKDPAPKSSKVDANEVEAL 300
QY 301 IVSPQKKNPPLFKULYKTFGPFLMSFFPKAIDHLMSSGPOILKLLKFNVDITAPD 360
DB 301 IVSPQKKNPPLFKULYKTFGPFLMSFFPKAIDHLMSSGPOILKLLKFNVDITAPD 360
QY 361 WQGFYTVLLFPVACLOTVLVHOYFHCFSQGRKIKTAVGAVYRKALVITNSARKSSTV 420
DB 361 WQGFYTVLLFPVACLOTVLVHOYFHCFSQGRKIKTAVGAVYRKALVITNSARKSSTV 420
QY 421 GEIVNLSVDAQRFMDLATYINMISAPLOVILLALYLLMLNLGSLVAGVAVWVLPVN 480
DB 421 GEIVNLSVDAQRFMDLATYINMISAPLOVILLALYLLMLNLGSLVAGVAVWVLPVN 480
QY 481 AVNAAKTKTYQVAMKSKDKRIRKLMNIIINGIKYLKIYANIELAKOKVLAIRQBELVYLK 540
DB 481 AVNAAKTKTYQVAMKSKDKRIRKLMNIIINGIKYLKIYANIELAKOKVLAIRQBELVYLK 540
QY 541 KSAVLAVGFTFWCTPEFLVALCTFAYVYVITIDENNIIIDAQTAFFSLAFNLIRPPLNLP 600
DB 541 KSAVLAVGFTFWCTPEFLVALCTFAYVYVITIDENNIIIDAQTAFFSLAFNLIRPPLNLP 600
QY 601 MVISSIVQASVSLKRLRIFLSHELBDSIERRPVKDGGTNSITVENATFTWARSDEPT 660

Db 601 MVISSIVQASVSLKRLRIFLSHEELEPDSIERRPVKOGGJNSITVRNATFTWARSDEPT 660
 QY 661 LINGITFSIPGALVAVVGVGCGKSSLSLALLAEMDKVEGHVATKGSVAVYPOQAMTOND 720
 Db 661 LINGITFSIPGALVAVVGVGCGKSSLSLALLAEMDKVEGHVATKGSVAVYPOQAMTOND 720
 QY 721 SLRENILFGCOLLEPPYRSVIQACALLPDLEILPSGDRTEIGEGKVNLSGGOKORVSLAR 780
 Db 721 SLRENILFGCOLLEPPYRSVIQACALLPDLEILPSGDRTEIGEGKVNLSGGOKORVSLAR 780
 QY 781 AVYGNADIYFLFDDPLSAVDHAVGHIFENVIQPKMLKNKTRILVTHSMSYLPQVDVYIV 840
 Db 781 AVYGNADIYFLFDDPLSAVDHAVGHIFENVIQPKMLKNKTRILVTHSMSYLPQVDVYIV 840
 QY 841 MSGKISEMGSYOELLARDGAFABFLRTYASTEODAEENGVTGVSQPKAKOMENG 900
 Db 841 MSGKISEMGSYOELLARDGAFABFLRTYASTEODAEENGVTGVSQPKAKOMENG 900
 QY 901 LVTDGAKOLORQSSSSSYSGDISRHHNSTAELOKAEKKEETMKLEADKAOTGOVKL 960
 Db 901 LVTDGAKOLORQSSSSSYSGDISRHHNSTAELOKAEKKEETMKLEADKAOTGOVKL 960
 QY 961 SVYWDYMKAIGLFISFLSIFLMCNHVSALASNYWLSMTDPIVNGTOEHTKVRLSYG 1020
 Db 961 SVYWDYMKAIGLFISFLSIFLMCNHVSALASNYWLSMTDPIVNGTOEHTKVRLSYG 1020
 QY 1021 ALGISOGIAVVGYSMAVSIIGIILASRCLHVDLHSLRSFMSFERTPSGNLVNRFSEKL 1080
 Db 1021 ALGISOGIAVVGYSMAVSIIGIILASRCLHVDLHSLRSFMSFERTPSGNLVNRFSEKL 1080
 QY 1081 DTVDMSIPEVIKMGSLFENVIGACIVTILATPIAIIIPPLGIYFFVQRFYVASSROL 1140
 Db 1081 DTVDMSIPEVIKMGSLFENVIGACIVTILATPIAIIIPPLGIYFFVQRFYVASSROL 1140
 QY 1141 KRLESVRSPPYSHFNFTLLGVSIVIRAFEBQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
 Db 1141 KRLESVRSPPYSHFNFTLLGVSIVIRAFEBQERFIHOSDLKVDENOKAYPSIVANRWLA 1200
 QY 1201 VRLECVNCIYLPALPAVVISRHSLSAGLVGLSVYSILOVTTYLNMLVRMSSEMETNIVA 1260
 Db 1201 VRLECVNCIYLPALPAVVISRHSLSAGLVGLSVYSILOVTTYLNMLVRMSSEMETNIVA 1260
 QY 1261 VERLKEVSETEKEAPMOIOETAPPSWPOVGRVFRNYCLRYREDLDPVLRHINVTING 1320
 Db 1261 VERLKEVSETEKEAPMOIOETAPPSWPOVGRVFRNYCLRYREDLDPVLRHINVTING 1320
 QY 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKITIIPQDPVL 1380
 Db 1321 EKVGIVRTGAGKSSLTGLFRINESAGEIIIDGINIAKIGLHDLRFKITIIPQDPVL 1380
 QY 1381 SGSLRNMULDPPSOYSDEEWTSLLELAHKDFVSALPKLDHECAEGEENLSVGORQLVCL 1440
 Db 1381 SGSLRNMULDPPSOYSDEEWTSLLELAHKDFVSALPKLDHECAEGEENLSVGORQLVCL 1440
 QY 1441 ARALLRKTILVDEATAVADLETDDLIOSTIRTOFEDCVLTIAHRLNTIMDYTRIVYL 1500
 Db 1441 ARALLRKTILVDEATAVADLETDDLIOSTIRTOFEDCVLTIAHRLNTIMDYTRIVYL 1500
 QY 1501 DKGEIOEYGAPSDILOQRGLFYSMAXDAGLY 1531
 Db 1501 DKGEIOEYGAPSDILOQRGLFYSMAXDAGLY 1531

Search completed: December 15, 2005, 15:17:29
 Job time : 185.197 secs

This Page Blank (upto)

This Page Blank (upto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:57:04 ; Search time 38.2211 Seconds
(without alignments)
4901.320 Million cell updates/sec

Title: US-10-665-283-8
Perfect score: 10016
Sequence: 1 MALRGFCSADGSDPLMDMNV.....RSVAVAKAKRFSIPDSLS 1947

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 9621673 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7849	78.4	1531	1 DVHUR	multidrug resistanc
2	4484.5	44.8	1527	2 JEO336	canalicular multist
3	3587	35.8	1545	1 S71841	multidrug resistanc
4	3507.5	35.0	1541	1 S71839	canalicular multist
5	3289	32.8	1494	2 E89447	protein F57C12.4 l
6	3225	32.2	1573	2 T21219	hypothetical prote
7	3195.5	31.9	1502	2 T42216	multidrug resistanc
8	2838	28.3	1515	1 S51863	cadmium resistance
9	2732	27.3	1398	2 T20434	hypothetical prote
10	2548	25.4	1478	2 T38712	ABC transporter SP
11	2384.5	23.8	1623	2 T01369	ABC transporter At
12	2356.5	23.5	1622	2 D86428	glutathione S-conj
13	2326	23.2	1495	2 E86428	probable ABC trans
14	2312	23.1	1444	2 T27408	hypothetical prote
15	2294.5	22.9	1559	1 S64757	probable membrane
16	2263.5	22.6	1488	2 F86428	glutathione ABC trans
17	2238.5	22.3	1516	2 F84919	glutathione-conjug
18	2230	22.3	1539	2 T48059	ABC transporter-11
19	2144.5	21.4	1355	2 T00961	hypothetical prote
20	2144.5	21.4	1514	2 T52080	multi resistance p
21	2140	21.4	1515	2 T52081	MMP-like ABC trans
22	2131.5	21.3	1490	2 T47840	multi resistance p
23	2129	21.3	1545	2 T46645	hypothetical prote
24	2123	21.2	1153	2 T26883	hypothetical prote
25	2109	21.1	1545	2 T42751	sulfonylurea recep
26	2104.5	21.0	1511	2 T42711	sulfonylurea recep
27	2098	20.9	1546	2 T42728	sulfonylurea recep
28	2069.5	20.7	1389	2 T47796	ABC transporter-11
29	2062.5	20.6	1661	2 S64800	probable membrane

30	2037	20.3	1121	2 C87973	protein Y43F8C.12
31	2014.5	20.1	1582	2 A56248	sulfonylurea recep
32	2008	20.0	1592	2 S48933	probable transport
33	1977	19.7	390	2 S68403	inward rectifier p
34	1962	19.6	390	2 JC4689	inwardly rectifyin
35	1905	19.0	390	2 A57616	inward rectifier K
36	1902	19.0	390	2 JC7901	inwardly-rectifyin
37	1888	18.8	1477	2 S64616	YOR1 protein - yea
38	1884.5	18.8	1037	2 T50518	ABC transporter-11
39	1805	18.0	1146	2 F84487	probable ABC trans
40	1722	17.2	946	1 JC5667	multidrug resistanc
41	1693	16.9	1548	1 DVINS	multidrug resistanc
42	1687	16.8	1421	2 T34225	hypothetical prote
43	1606	16.0	1427	2 T20903	hypothetical prote
44	1575.5	15.7	1427	2 T39219	acp-binding caset
45	1558	15.6	1469	2 T50210	probable ABC trans

ALIGNMENTS

RESULT 1

DVHUR

N:Alternate names: multidrug resistance-associated protein (MRP)

C:Species: Homo sapiens (man)
C:Date: 30-Jun-1993 #sequence_revision 05-Dec-1998 #text_change 19-Jan-2001

C:Accession: A44231; A37495

R:Coile, S.P.C.; Bhargava, G.; Gerlach, J.H.; Mackie, J.E.; Grant, C.E.; Almqvist, K.C.; Science 258, 1650-1654, 1992

A:Title: Overexpression of a transporter gene in a multidrug-resistant human lung cancer cell line

A:Reference number: A44231; MUID:93088080; PMID:1360704

A:Accession: A44231

A:Molecule type: mRNA

A:Status: nucleic acid sequence not shown

A:Residues: 1613-1623

A:Cross-references: UNIPARC:UPI00001746C; GB:L05628; NID:G1835658

A:Experimental source: small cell lung carcinoma cell line H69AR

A:Note: Sequence extracted from NCBI backbone (NCBI:P.119851); this sequence has been corrected.

R:Coile, S.P.C.; Deele, R.G. Science 260, 879, 1993

A:Title: Multidrug resistance-associated protein: sequence correction.

A:Reference number: A37495; MUID:93262415; PMID:8098549

A:Accession: A37495

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-60 <CO2>

A:Cross-references: UNIPARC:UPI00001746C; GB:L05628; NID:G1835658

A:Note: sequence extracted from NCBI backbone (NCBI:P.131929)

C:Genetics:

A:Gene: GDB:MRP

A:Cross-references: GDB:136335; OMIM:158343

C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

C:Keywords: antibiotic resistance; ATP; duplication; nucleotide binding; P-loop; transmembrane

F:678-685/Region: nucleotide-binding motif A (P-loop)
F:788-792/Region: nucleotide-binding motif B
F:110-150/Domain: ATP-binding cassette homology <ABC2>
F:127-133/Region: nucleotide-binding motif A (P-loop)
F:150-145/Region: nucleotide-binding motif B

Query Match 78.4%; Score 7849; DB 1; Length 1531;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1529; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy	1	MALRGFCSADGSDPLMDMNVNTNTSNDPFCQNTLVWVPCVLAACPFYLYSRH	60
Db	1	MALRGFCSADGSDPLMDMNVNTNTSNDPFCQNTLVWVPCVLAACPFYLYSRH	60
Oy	61	DRGYQMTPLNKTALGFLMTIVCWADLFYFWMERSGIFLAPVFLVPTLGIITLLA	120
Db	61	DRGYQMTPLNKTALGFLMTIVCWADLFYFWMERSGIFLAPVFLVPTLGIITLLA	120

```
QY 121 TFLIQLERRKGVGGSGIMLTFWLVAALCALILRSKIMTALKEDAQVDLFEDITFYFYS 180
D 121 TFLIQLERRKGVGGSGIMLTFWLVAALCALILRSKIMTALKEDAQVDLFEDITFYFYS 180
QY 181 LLLIQLVLSCSDSPISFSETIHDNPPCESSASLSTITFMWITGLIVRYRQPLESSD 240
D 181 LLLIQLVLSCSDSPISFSETIHDNPPCESSASLSTITFMWITGLIVRYRQPLESSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVKVYSSKDPAPCKESSKYDANEVEAL 300
D 241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVKVYSSKDPAPCKESSKYDANEVEAL 300
QY 301 IVKSPOKEWNPDLFKVLYKTFGPYFLMSFFPKAIHDLMMFSGPOLIKLIFVNDTKAPD 360
D 301 IVKSPOKEWNPDLFKVLYKTFGPYFLMSFFPKAIHDLMMFSGPOLIKLIFVNDTKAPD 360
QY 361 MQGFYFVVLVPTVCTCLOTLVLAHQYFHI CFVSGMRITKZAVIGAVYRKALVITNSARKSTV 420
D 361 MQGFYFVVLVPTVCTCLOTLVLAHQYFHI CFVSGMRITKZAVIGAVYRKALVITNSARKSTV 420
QY 421 GEIYNLMSVDAQREMDLATYINMIWSAPLOVILALYLLMLNLSBSVLGAVAVMVLMPVN 480
D 421 GEIYNLMSVDAQREMDLATYINMIWSAPLOVILALYLLMLNLSBSVLGAVAVMVLMPVN 480
QY 481 AVMAAMKTKTYOVAMHKSNDRIKLMNBIINGIKVILKIYAMELAFKDKVLAIROEBLKYK 540
D 481 AVMAAMKTKTYOVAMHKSNDRIKLMNBIINGIKVILKIYAMELAFKDKVLAIROEBLKYK 540
QY 541 KSAVLASVGTFTWCTPPLVALCTPAVAVVTIDENNIILDAQAPVSLALFNLRPLNLTLP 600
D 541 KSAVLASVGTFTWCTPPLVALCTPAVAVVTIDENNIILDAQAPVSLALFNLRPLNLTLP 600
QY 601 MVISIYQASVSLKRLRIFLSHEBELPDSERRPVKDGCGNSTVRNATFTMARSDPT 660
D 601 MVISIYQASVSLKRLRIFLSHEBELPDSERRPVKDGCGNSTVRNATFTMARSDPT 660
QY 661 LMGITFSIPBEGALVAVVQVCGKSLISALLAEMDKVEGHVAIKGSVAVYVPOQAMIOND 720
D 661 LMGITFSIPBEGALVAVVQVCGKSLISALLAEMDKVEGHVAIKGSVAVYVPOQAMIOND 720
QY 721 SLRNNILFGCOLLEPYRYSVIOACALLPDLBELPSGDTEIEGKCVNLSGGQKRVSLAR 780
D 721 SLRNNILFGCOLLEPYRYSVIOACALLPDLBELPSGDTEIEGKCVNLSGGQKRVSLAR 780
QY 781 AVYENADITLFPDDPLSAVDAHVGKHI FENVI GPKMKNKTRILVTHSMSTLPQVDVITV 840
D 781 AVYENADITLFPDDPLSAVDAHVGKHI FENVI GPKMKNKTRILVTHSMSTLPQVDVITV 840
QY 841 MSGKISMSGSYOELLARDGAPAEFLRTYASTEOBDAENGVTGVSQPGKEAKOMENGM 900
D 841 MSGKISMSGSYOELLARDGAPAEFLRTYASTEOBDAENGVTGVSQPGKEAKOMENGM 900
QY 901 LVTPSAGQOLOROLSSSSSYSGDISRHHNSTAELOKAAKKEBTMKLEADKAQOGYKL 960
D 901 LVTPSAGQOLOROLSSSSSYSGDISRHHNSTAELOKAAKKEBTMKLEADKAQOGYKL 960
QY 961 SVYDYMKAIGLFI SFSLIFLPMCHVSAALSNVWLSLMTDPTINGTQHTKVRLSYVG 1020
D 961 SVYDYMKAIGLFI SFSLIFLPMCHVSAALSNVWLSLMTDPTINGTQHTKVRLSYVG 1020
QY 1021 ALGISOGIAVEGSAVAVSIGGILASRCLHVDLHLSILRSFMSFFERTSPGULVNFSEKL 1080
D 1021 ALGISOGIAVEGSAVAVSIGGILASRCLHVDLHLSILRSFMSFFERTSPGULVNFSEKL 1080
QY 1081 DTVDSMTPEVITKMGMSLPNVI GACTVILLATPIAIIIPGLGIYFVQRPYVASRQL 1140
D 1081 DTVDSMTPEVITKMGMSLPNVI GACTVILLATPIAIIIPGLGIYFVQRPYVASRQL 1140
QY 1141 KRLSVSSPVYSHNEFTLLGVSVIRAFEEQERFTHOSDKVDENOKAYYSIVANRWLA 1200
D 1141 KRLSVSSPVYSHNEFTLLGVSVIRAFEEQERFTHOSDKVDENOKAYYSIVANRWLA 1200
```

```
QY 1201 VRLCEVNCIVLPALPAVVISRHSLSAGLYGSLYSLOQVTTYINMLVRMSSEMETNIVA 1260
D 1201 VRLCEVNCIVLPALPAVVISRHSLSAGLYGSLYSLOQVTTYINMLVRMSSEMETNIVA 1260
QY 1261 VERLKEYSETEKEAPMOIOETAPPSMPQYGRVFRNYCIRYREDLPVLRIHINVTINGG 1320
D 1261 VERLKEYSETEKEAPMOIOETAPPSMPQYGRVFRNYCIRYREDLPVLRIHINVTINGG 1320
QY 1321 EKVGIVRGTGAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
D 1321 EKVGIVRGTGAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVLF 1380
QY 1381 SGSLRNMULDPSQYSDSEWNTSLBLAHKDFVSAALPKLDHECAGGENLSVGGROVLCL 1440
D 1381 SGSLRNMULDPSQYSDSEWNTSLBLAHKDFVSAALPKLDHECAGGENLSVGGROVLCL 1440
QY 1441 ABALIRKTKILVDEATRAVDLETDDLIQSTIRIQFEDCVLTIAHRLNTIMDYTRVYL 1500
D 1441 ABALIRKTKILVDEATRAVDLETDDLIQSTIRIQFEDCVLTIAHRLNTIMDYTRVYL 1500
QY 1501 DKGEIOEYGAPSDLLQQRGLFYSMAKXAGLY 1531
D 1501 DKGEIOEYGAPSDLLQQRGLFYSMAKXAGLY 1531

RESULT 2
JE0336
canalicular multispecific organic anion transporter - human
C:Species: Homo sapiens (man)
C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004
C:Accession: JE0336
R:Juchimul, T.; Hinojosa, E.; Haga, S.; Nakamura, T.; Tanaka, T.; Tsh, S.; Furukawa, M.;
Biochem. Biophys. Res. Commun. 252, 103-110, 1998
A:Title: Isolation of a novel human canalicular multispecific organic anion transporter,
t.
A:Reference number: JE0336, MUID:99032812; PMID:9813153
A:Accession: JE0336
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-1527 <UCH>
A:Cross-references: UNIPROT:O15438; UNIPARC:UPI0000169859; GB:AF083552
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C:Keywords: ATP
F:11306-1499/Domain: ATP-binding cassette homology <ABC2>

Query Match 44.8%; Score 4484.5; DB 2; Length 1527;
Best Local Similarity 56.6%; Pred. No. 1.5e-294;
Matches 871; Conservative 261; Mismatches 372; Indels 35; Gaps 9;

QY 8 SAGSDPLMDNVTWNTSNPDFTYCFONTYLVWVPCFYLMACFPFYLYLSRHDGTYIQM 67
D 8 SAGSDPLMDNVTWNTSNPDFTYCFONTYLVWVPCFYLMACFPFYLYLSRHDGTYIQM 67
QY 7 SGEIGSKFWDNLSVHTENPDLTPCFQNSILLAWPCILYLVWALPCYLLYLRHHGCGYITL 66
D 7 SGEIGSKFWDNLSVHTENPDLTPCFQNSILLAWPCILYLVWALPCYLLYLRHHGCGYITL 66
QY 68 TPLAKTKALGFLMIYICMADLFYSFWERSGIFLAVPVLSPTLLGITTLATPILIOLE 127
D 68 TPLAKTKALGFLMIYICMADLFYSFWERSGIFLAVPVLSPTLLGITTLATPILIOLE 127
QY 126 7 SHLSKLMVGLVLLMCVSWMDLFYSFGLVGRAPAPVFTPTPLVVGVTMLLATLLIOYE 126
D 126 7 SHLSKLMVGLVLLMCVSWMDLFYSFGLVGRAPAPVFTPTPLVVGVTMLLATLLIOYE 126
QY 128 RRGVQSSGIMLTFWLVAALCALILRSKIMTALKEDAQVDLPFDITFYVFSLLIOQV 187
D 128 RRGVQSSGIMLTFWLVAALCALILRSKIMTALKEDAQVDLPFDITFYVFSLLIOQV 187
QY 187 RLGQVQSSGIMLTFWLVAALCALILRSKIMTALKEDAQVDLPFDITFYVFSLLIOQV 186
D 187 RLGQVQSSGIMLTFWLVAALCALILRSKIMTALKEDAQVDLPFDITFYVFSLLIOQV 186
QY 188 LSCFSDSPISFSETIHDNPPCESSASFLSITFMWITGLIVRGYROPLESDLSLNKE 247
D 188 LSCFSDSPISFSETIHDNPPCESSASFLSITFMWITGLIVRGYROPLESDLSLNKE 247
QY 247 LACFRKRPFPFSACNVDPNPFERISAGFLSLFLFWMTFKMAIYGRHPLLEKDLMSLKEE 246
D 247 LACFRKRPFPFSACNVDPNPFERISAGFLSLFLFWMTFKMAIYGRHPLLEKDLMSLKEE 246
QY 248 DTSQVVPVLVKNMKKECAKTRKQPVKVYSSKDPAPCKESSKYDANEVEALIVKSPQK 307
D 248 DTSQVVPVLVKNMKKECAKTRKQPVKVYSSKDPAPCKESSKYDANEVEALIVKSPQK 307
QY 307 DRSQVVPVLVKNMKKECAKTRKQPVKVYSSKDPAPCKESSKYDANEVEALIVKSPQK 294
D 307 DRSQVVPVLVKNMKKECAKTRKQPVKVYSSKDPAPCKESSKYDANEVEALIVKSPQK 294
QY 308 EMNLSLFRVLYKTFGPYFLMSFFPKAIHDLMMFSGPOLIKLIFVNDTKAPDMQGYRYT 367
D 308 EMNLSLFRVLYKTFGPYFLMSFFPKAIHDLMMFSGPOLIKLIFVNDTKAPDMQGYRYT 367
QY 367 R-KPSFLKALATGSSFLISACFKLIQDLISLNPOLSTILIRPISNMPGSPMWGFLVA 353
D 367 R-KPSFLKALATGSSFLISACFKLIQDLISLNPOLSTILIRPISNMPGSPMWGFLVA 353
```

QY	368	ULLPFLACLOTVLHVOYHICFVSQMBRIKTVANGVVRKAVITVMSAKRSVGEIVNIM	427
Db	354	GLMFCSNMQSLLOHYYHYTFVGVKFRPTGIMGYIRKALVTINSVVRASVGEIVNIM	413
QY	428	SVDAORFMDLTAIYINMIMSAPLOVILATYLMLNLTNGSPSLAGVAVNVLVMPVNAVMANKT	487
Db	414	SYDAORFMDLAPFNLMLMSAPLOIILATILYFLMQNIGPSVLGAAMVLLIPLNGAVAKM	473
QY	488	KTYOVANHMSKONRIKLMNEILINGIKVLKLYAMBLAFKDKVLAIROEBIKVLKKSAYLSA	547
Db	474	RAFYQYOMKLIKOSRIKLMSEILINGIKVLKLYAMEDSPFKQVGVIRGQEBLOLRTAVALHT	533
QY	548	VGTFFWTCCTPFLVALCTFAVYVTTIDENNILDAQTFVSLALPNLRPLNLTIPWYISSTV	607
Db	534	TTTFPMWCSPLVLTITLTMVYVYVDPNNVLDAEKAFVSYSLEPNIRLPLNMLPOLISNLT	593
QY	608	QASVSLKRLIRPLSHEELEPDSIERBPVYDGGGTNSITVRNATPTWASDPPTLNGIFNS	667
Db	594	QASVSLKRIQOPLSBEELDPOGVEKTTISPG--YAITIHSTFTWAOULPPTLHSLIQ	650
QY	668	IPEGALVAVVGQVGGKSSLSALLAENDKXVEGHVAIKGSVAYVPOQAMIQNDSIRENLT	727
Db	651	VPKGAALVAVGPGVGGKSSLSVALLGEMKELEGKVMKMGSVAYVPOQAMIQNCTIQENVL	710
QY	728	FGCQLEBYYSVIOACALLPDLLETLPESGDRTEIEBKVNISGGQKORVSARAVYSNAD	787
Db	711	FGKALNPKRYOOLTEAALCLADLEMLPGGDQTEIEKEGINISGGQKORQVSLARAYSPAD	770
QY	788	IYLPDPLSADVAHNGKHIFENVIGPKMLKNKTRILVTMSYVPOVDVITVW6GGKIS	847
Db	771	IFLDDPLSAVDSHAKHIFDHVIGPEGLAKGTRVLVTHGSLFSPQDFTFIVLADGOVS	830
QY	848	EMGSYOELLARDGAFAEBELRTYASTBOEDAEENGVTGVSQPKKAKOMENGML-----	901
Db	831	EMGYPALALQNGNSPANFLCNVAPBEDDGHLEDSYTALEGAEDKALLIEDTLNSHTLUT	890
QY	902	----VTDSAGQOLQOLSSSSSSYSGD-----ISRHNSTAE-LQKBAKKEBYTKLMEAD	951
Db	891	DNDPVTYVVOKQPMKOLALS-DEGOGGRPVPRHRLGPSEKQVYTKAKADGA--LTQEB	947
QY	952	KAQNGQVLESVYMDWKRIIGLFIPLSLFLEPCHNVSLASNMYLSMTDDPIYNGTQEH	1011
Db	948	KKAIQTVELSPFMDAKAVGLCTTLTALICLLVYGQAAIIGAVNWLISANTNDAMASRONN	1007
QY	1012	TKVRLSVYGALGISQGIAPFGSMAVNSIGGILASRLCHVDLHSLTRSPMSEFFERTPSGN	1071
Db	1008	TSRLRGVYALGIILOGFLVMLAAMMAAGGIOARVLMQALLHNRKIRBPQSFDTTPSGR	1066
QY	1072	LVNRFSEKLDIVDSMIPEVIKMFMSLFPNVIIGACIVILLATPIAIIIPRLGIYFFVOY	1131
Db	1068	IINCFSKOIYVYDDEVILAVIIMLTINSFNAISTLVIVMASIPLFTVILPLAVIYTLVQY	1127
QY	1132	FYVASSRDLKLEBSRSRPVYSHMETLIGSVIRAFEOERFIHQSLKVDENOKAYYP	1191
Db	1128	FYAAISROLKLEBSRSRPIYSHFSEYTVGASVIRAVYRSBDFEIIPTKVDANORSCP	1187
QY	1192	SIYVARNALVBLECYGNCIVFAALFAVYSRHSAGLVGVSYSLSLQVTTYLMNLVMS	1251
Db	1188	YIISKRWLSIGVEFPNGCVLPLPALFAVYIGSSLWPLGVLAGSVSISLOVYTRALNMIMMM	1247
QY	1252	SEMETNIVAVERLKEYSETEKBAPOQIOETAPPSMPOVGVFEFANYCIARREIDLFPYLR	1311
Db	1248	SDLESNIVAVERVKEYSKTEPTEAPVWVGSNRPEGMPPRGVEFEPFANSVVRVRPGLDVLR	1307
QY	1312	HINVTINGEKYVIGRTGAGKSSITLGLFRINSEABEIIIDGINIAKIGLHDLRFKIT	1371
Db	1308	DLSLHVGHEKYGIVGRTGAGKSSMTLCLFRLENAKEIRIDIGLVNADIGHADVRSQLT	1367
QY	1372	IIPDPVYFSGSLRNLNDPFSQYSDPEWTSLEHLNLDOPYSALPDKLDHCAEGEYMS	1431
Db	1368	IIPDPILFSGTLRNNDPFGSYSEBDIWMALLESHTLTPVSSQAPAGLDPOCSBEGEYMS	1427
QY	1432	VGQROVLCLARALLKTKILVLDXYAAVLDLTDLLQSTIRTOFEDCTVITLHARLNTI	1491

Db	1428	VGQROLVCLARLLARKSILVLDESTPAIDLETDLVQATIRTPQDPTCTVLTIAIRLNTI	1487
Qy	1492	MDYTRVIVLDKGEIOEYGA PSDLLQOGRGLFYSMANDAGL	1530
Db	1488	MDYTRVIVLDKGVVAEFPDPSANLTIARGIETFGMARDAGL	1526
RESULT 3			
71841			
multidrug resistance protein, canalicular - human			
C.Species: Homo sapiens (man)			
C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001			
C.Accession: S71841; S71840			
R.Koenig, J.; Keppler, D.			
submitted to the EMBL Data Library, August 1996			
A.Reference number: S71841			
A.Accession: S71841			
A.Molecule type: mRNA			
A.Residues: 1-1545 <KOE>			
A.Cross-references: UNIPARC:UPI00001746CD; EMBL:X96395; NID:g1507819; PIDN:CA65259.1;			
R.Buechler, M.; Koenig, J.; Brom, M.; Kartembeck, J.; Sprinz, H.; Horle, T.; Keppler, D.			
J. Biol. Chem. 271, 15091-15098, 1996			
A.Title: cDNA cloning of the hepatocyte canalicular isoform of the multidrug resistance			
A.Reference number: S71839; MUID:96279006; PMID:8662992			
A.Accession: S71840			
A.Status: nucleic acid sequence not shown			
A.Molecule type: mRNA			
A.Residues: 1415-1429, 'VP', 1432-1455, 'E', 1457-1545 <BUE>			
A.Cross-references: UNIPARC:UPI00001746CE; EMBL:X96395			
C.Genetics:			
A.Gene: GDB:ABCC2; CMOAT; ABC; MRP2; CMRP; DS			
A.Cross-references: GDB:6089489; OMIM:601107			
A.Map position: 10q24-10q24			
C.Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology			
C.Keywords: ATP; glycoprotein; nucleotide binding; P-loop; transmembrane protein			
F:654-837/Domain: ATP-binding cassette homology <ABCL>			
F:671-678/Region: nucleotide-binding motif A (P-loop)			
F:1137-1510/Domain: ATP-binding cassette homology <ABCS>			
F:1134-1341/Region: nucleotide-binding motif A (P-loop)			
Query Match			
Beat Local Similarity 47.5%; Score 3587; DB 1; Length 1545;			
Matches 746; Conservative 286; Mismatches 461; Indels 78; Gaps 20;			
Qy	3	LRGCSADGSDPLMDMNTVNTS-----NDPTKCFONTLVWPCFYLMACFPYFL--	55
Db	2	LEKFCN-----STFNMSFLDSEADLPICFEQTVLWVPIGLGMLLAPQQLHV	51
Qy	56	YLSHHDRCYIQMTPLANKTKTALGFLMWVCADLFYSWESRSGLFLAPVPLVSTTLGI	115
Db	52	YKSRKRSSTTKLYAK-QVFVGFL-LLAALVALVLTESGQATVAVYTPSLY-L	108
Qy	116	TTLATFPIQLERRKGVSSGIMLT-FWLVALVCALILRSKIMTALKEDAQVDLFRDIT	174
Db	109	GTWLLVLLIYQSRQMCQYKNSWFLSLFWILSLICGTPOQLIRLITLQGD-NSNLAYCCL	167
Qy	175	FYVYFSLILVLVSCFDSRSLPFSETHDNPCESSASFLSRITFWITGLIVGRQ	234
Db	168	FPISGFIILLIFSAFSEN-----NESSNPSSIASFLSSITYSWYSLILKGYKR	219
Qy	235	PLBSGDLMSLKEPTSEBQVAVLVKMMKECAKTRK-----QVVKVYSSKD	281
Db	220	PLTIEDVWEVDEEKKTKLVSKFETHMKRELQKRRALQRROEKSSOONSQARLGLKXN	279
Qy	282	PAQPSKVDANEVEALIVKSPQKWNPS--LFKVLKTFGEYFLMSFFPKAIHIDLM	339
Db	280	QSQSDALVL--DVEVERKKKSGTKKDVPSKMWLAKLFKTFVWLLKSLFLKLVDIET	336
Qy	340	FSGPOLIKLILKFNDRTPAKPWQGYFYTVLLEFVACIQTLVLIHQYFIHCPVSGRIKAV	399
Db	337	FVSPQLKLLISFASDRDTYIMIGYLCALILFTALILQSFCLOCFDLGCFVLGKVKRAI	396

Qy	400	IGAVRYKRLVITNSRKSTYGEIYNLMSVDAQRMDLATYINMISAPLOVITLALYLLM	459
Dp	397	MASVYKKALTTLSNLRKKEYTVEGETYANLMSVDAQKMDVYFNHMLMSVLOIVLISIFLW	456
Qy	460	LNLGYSVLAGVAVWVIAWYVAVNAVMAMKCTYQVVAHKSQNRKIKLMNTEIINQIKYLKUYA	519
Dp	457	RELGSVLAGVAVWVLPINAILSTKSTIOVKRMOKDKRLKIMNELSGIKLKYKPA	516
Qy	520	WELAFKQVLAIRQBELKVLKKSAVLSAVGTFTWCTPFLVALCTFAVYVITDENNIIDA	579
Dp	517	WEPSPRDQVQMLRKKELEKNILAFSLOQCQVLIFFVQFOTPLVAVSVTFVSYYVLVDNSNIIDA	576
Qy	580	QTAFVSLALENLRPELNLIPWVSISSYQASYSLKRLRIFLSHELEBPSIRRPVYXOG	639
Dp	577	QKAFPSITLFLNLRPELPSLPMWISSMLQASVSTERLEKYGDDLDTSAIRHSCNFD--	634
Qy	640	GTNSITVNAFPTWARSDBPTLNGITFSPBECALVAWVQVQCGSSLSLALAMDKYE	699
Dp	635	--KAMQSEASFTWHDSEATYRDVNDIMAGQVAIVGPVQSGSSLSLSALGEMENYH	692
Qy	700	GHVALKGSVAAYVPOQAMIQNSDLARENILFGCOLBEPYRSVIOACALLPDLIELPSGRT	759
Dp	693	GHITTKGTTAYVPOQSMWINGTIDKONILFGEFNEKRYQVULAEACALLPDLIELPGDILA	752
Qy	760	EIEGEGVNLSSGQOKRVSLARVYSNADITYLEPDDL SAVDAVGHIPENYVGPYKMLXN	819
Dp	753	EIEGEGVNLSSGQOKRISILARTYONLDIYLLDDPLSAVDADVGHIFPKVYGPGLLKG	812
Qy	820	KTRLIVTHSMKVLPOVDYIVMWSGKISBMGSYOELLRDGAFAFELRTY-ASTROEODA	878
Dp	813	KTRLIVTHSMKFLPOVDEIVLGNQITVEKGSYSALLAKKGFAKOLKTFLLHRTPEBEA	872
Qy	879	EENGVTVSGPKEKQOM-----ENGMVLTDSAGKOLQORULSSSSSYSGDISRHH	928
Dp	873	-----TVHDSGHEEADYGLISSVEIPEDASITMRRENSPRRTLSSSSRNGHKLMSL	927
Qy	929	NSTALEQAEAKKET-----WKMEADKAQCYQYKLSYWWYMYKALGLIFLSLIFLM	983
Dp	928	RNSLKTRVNSLSKEDEBELVKQKLLKEKEIEFGKVKFSIYEYVLOAIGLFSIFPILIAFV	987
Qy	984	CNHVSALASNYLSMT--DDPVTNGQ--BHTYKRLSYGLAGISOGIANGYSMAVSI	1033
Dp	988	MNSVAFISNMLMSMTSDSKIFNSTDYIPASGRDMKVGYYGLAGLQGI FVFI AHFWSAF	1044
Qy	1040	GGILASRCLHVDLHSLRSPMSFEERPPSGNLVNRFSKELDTVDMSIPEVIMMGSLF	1095
Dp	1048	GFVHASNILHKQILNNILRAPMRFPTDTPGRI VNRPAIGDISIVDQTLPOSLRTWITCGL	1107
Qy	1100	NVIGACVILIATPLAAILIIPPLGLIYFPVQAFYYAASSQOLKLESVSRSRPYSHFNETL	1155
Dp	1108	GIISTLVICMATPFTIIVIPGLIITYSVQMFYVSTSRQILRIDSVTRSPYISHFSEAV	1167
Qy	1160	LGVSVIYIRAFEEORERTIHOSDLKVDENOKRAYYSIYANRMLAVRLCVCNGCIVLPAALFAY	1211
Dp	1168	SGLPVIRAFEHQORILKKNRVARIDNOKCVSWITSNKLARLRLVGNLIVFSBALMV	1222
Qy	1220	ISRHSLSAGLVGLSVYSLSQVTTYLNMLVRMSSEMETNIVAVERLEKEYSETEKESAPMOQI	1273
Dp	1228	IYRDTLSGDTYGFVLSNMLNITQTLNMLVYRMSSELETNIVANERTLEYTKVNEAPW-VT	1286
Qy	1280	ETAPSSSPQVQRYVREFRNYCLARYEDDDPVLRIHNVTTNGEKVKGI VQRTGAGKSLTIG	1333
Dp	1287	DKRPPDPMPSPGKGIQFNMYQVARYPELDVLVIRGICIDIGSMEKIIVVGRGTGAGKSLTNC	1340
Qy	1340	LFRINESAGEIILIIDGINIAKIGLHDAFKITIIPODPLPFGSIRMLMDPPSOYSDEBV	1397
Dp	1347	LFRILEAAGQIILIIDGVDIASIGLHDLKRLTIIPODPLTFSGSIRMLMDPPNNYSDEBI	1406
Qy	1400	WTSLELAHAKDFVSLPDKLDHECAGEGENISVGOQROI VCLARALLRKTIKILVLEATVA	1455
Dp	1407	WKALELAHLKSFVNSLOGLSHGHEFAGCNLSIGQROLLCLGRALLRKSKITIVLEABATRA	1466
Qy	1460	VDLETDDLIQSTIRTOFEDCTVLIYAHRLNTIMDYTRVIVLDKGEI QEYGAESDILLQORG	1513

```

Db      1467 VLEFDNLIQTIIQNSFACHTVITIAHRLHTIMDSKRWMLDNGKILIEGSPBEELQIPG 1526
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Qy      1520 LFYSMAKDAGL 1530
      ||:::|:::|
Db      1527 PFYFMAKEAGI 1537

RESULT 4
S71839
canaliculular multidrug resistance protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S71839
R:Buechler, M.; Koenig, J.; Brom, M.; Kartenbeck, J.; Spring, H.; Horie, T.; Keppler, D.
J. Biol. Chem. 271, 15091-15098, 1996
A>Title: cDNA cloning of the hepatocyte canalicular isoform of the multidrug resistance
A:Reference number: S71839; MID:96279006; PMID:8662992
A:Accession: S71839
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1541 <BDB>
A:Cross-references: UNIPROT:Q63120; UNIPARC:UPI000012578; EMBL:X96393; NID:G1292881; PI
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C:Keywords: ATP; glycoprotein; nucleotide binding; P-loop; transmembrane protein
F:120-124/Domain: transmembrane #status predicted <TM01>
F:127-151/Domain: transmembrane #status predicted <TM02>
F:160-187/Domain: transmembrane #status predicted <TM03>
F:305-329/Domain: transmembrane #status predicted <TM04>
F:354-381/Domain: transmembrane #status predicted <TM05>
F:431-451/Domain: transmembrane #status predicted <TM06>
F:456-476/Domain: transmembrane #status predicted <TM07>
F:536-564/Domain: transmembrane #status predicted <TM08>
F:574-602/Domain: transmembrane #status predicted <TM09>
F:650-833/Domain: ATP-binding cassette homology <ABC1>
F:667-674/Region: nucleotide-binding motif A (P-loop)
F:966-994/Domain: transmembrane #status predicted <TM10>
F:1018-1046/Domain: transmembrane #status predicted <TM11>
F:1104-1132/Domain: transmembrane #status predicted <TM12>
F:1203-1228/Domain: transmembrane #status predicted <TM13>
F:1331-1506/Domain: ATP-binding cassette homology <ABC2>
F:1330-1337/Region: nucleotide-binding motif A (P-loop)
F:1007,1010,1011/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match      35.0%; Score 3507.5; DB 1; Length 1541;
Best Local Similarity 46.2%; Pred. No. 2e-228;
Matches 721; Conservative 301; Mismatches 471; Indels 69; Gaps 21;

Qy      6 PCSADGSDPLMWNNTWNTSNDFPKCFQNTVLVWDCFLYMACPF--YFLYLSRHDRG 63
Db      4 FCNS----TFWPLSL-ESPEDLPICFEQTVLWITPLGLFIMLAPQLVSVRSRTKRS 58
Qy      64 YIQMTPLNKTALGSLIIVICWADLFYSFWESRSGFLAPVFLVSPITLGIITTLAFL 123
Db      59 SI--TKFYIAKQVFFVFLILIAIDISLALTEBTGQATVPVRVYTNP-ILVLCTWLVLVA 115
Qy      124 IQLERRKGVOSSGIMLT-FWLVAIVCALAIRSKIMALKEDAOVDLFRDITFYVFSLL 182
Db      116 VGHSGQMCYRKXKSWPLSLWILSVLCGVQPQF-LITALLKDSNNMAVSYLFFSYSGFQ 174
Qy      163 LIQLVLSCFSDRSPLEFSETHIDNCPRESSASFLSRITFMWITGLVGRYQPLSGSLW 242
Db      175 IYLLILITAFSGPS-----DSTQTPSVASFLSITFSWYDRTVLKGYKGPLTLBDW 226
Qy      243 SLNK-----EDTSEQVVPVLVKN-----WKKECAKTRKQPVKVVYS-SKPAQPKESK 290
Db      227 DIDEGFKTRSVTSKFEAATKTDLOKARQAFORLRQSKQRKEATLHGLNKKQSOQDVLV 286
Qy      291 VDANEVEVALIYVSPQKEMNPSLFKVLYKTPGQFPLMSPEPKAIIHDLMMFGSPQILKLLI 350
Db      287 LBEAKKSKSKTKTDYIKSM---LTKSLFKTFHVIILKSLFKLIDHLLVFLNPLQILKLLI 343
Qy      351 KRVNDTKADWQGYFYTVLLFVTACIQTLVLAHQYFHICFVSGMBRIKTAIVAGAYVRKALVI 410

```

```

Db      344 GFVSSNSYVWFYICAILMFAVTLTIOFCLQSYFOHCFVLGMCVTRTVSSYIKKALLTL 403
Qy      411 TNGARKSTVEIYVMSVADQRFMDLATYINMWSAPLOYITLALYMLNGSVLACV 470
Db      404 SNLARKYITGETVMSVDSQKLMDATNNMQLVMSVYIQTLLIFFLMRELGPSSILAGV 463
Qy      471 AVNVLAVPVNAVMAMKTKTYOVAMHKSQKNRIKLMNEIINGIKYLUKUYAMELAFKDKVLA 530
Db      464 GWVLLLPVNGVLTAKTRINIOVQNMKNKDKLIMNELLSGIKLTKTPAMEPSPQBOYQG 523
Qy      531 IROBELFLVKKSAVLSAVGFTWVCTPPLVALCTPAVYVITDENNIIIDAQTAFFVSLAFN 590
Db      524 IRKRELKMLRFQGLQSLIFILQITPLVSVTFESVYVLVDANVLAERAFSITLFFN 583
Qy      591 ILRPLMLPMVTSIYQASVSLKRLRIFLSHELEPDSIERRPYKDGCGTNSITVRNAT 650
Db      584 ILRPLMLPMVTSIYQASVSVRLERYLGGDDLTSAIRRVSNPD---KAVFSEAS 639
Qy      651 FTWASDPPTLNGITFSPREGALVAVVGVGCGKSSLSLALAMDKVEGHVATKGSVAY 710
Db      640 FTWDPDLBATIQDNLDIRKQOLAVVGVGSGKSSLSVAMLGEMENVGHITTOGSTAY 699
Qy      711 VPQAMTQNDLSRENIIFGCOLBEPYRSVIOACALLPDLIELPBGRTTEIGEGVNLG 770
Db      700 VPQSMVONGTIKNNILFGSEYNEKXYQOVAKACALLPDLIELGDMAEIGEGVNLG 759
Qy      771 GQKQVSLARAVSNADYILFDPLSAVDAAHVGHIFENYVIGPKMLKNKTRILLVTHSMS 830
Db      760 GQKQVSLARAAVYODADIYILDDPLSAVDAAHVGHIFENYVIGPKMLKNKTRILLVTHSIH 819
Qy      831 YLPQVDTIYVSGKISEMGSYOBELARDGAFAPFLKTYAS-TQEOBDAENGVTVGSGP 889
Db      820 FLRPVDIIVVAGKTKILEKSYRDLKKGVPARNWMTFMKSGPDEEBATVNN----- 872
Qy      890 GKEAKQENGMVLT-----DSAG-----KOLQROSSSSSYSGDISRHNNSTAEIQ-- 935
Db      873 DSEBDDDDGLIFPMEIIPEDASLARKRENSLRKTSRSRSRSSRSGKSLAKSLKTKV 932
Qy      936 ---KAEAKKETWKLMEADKAQTOGVLSVYWDYMKALGLEIFSLIFLPMCHVSLAS 992
Db      933 NVLKEKEVEGQGLIKKEFVETGKVFSYLYKTLQAVGWSLIFILFLFGVLANVAFISG 992
Qy      993 NYWLSLMT-DDPIYNGR---QEHKVALSVYGAIGISOGIAVFGVSAVSTGGILARCL 1048
Db      993 NLMWLSAITSDDNLNGTNNSSSHRDMRIGVFAGLGAQIGICLLISTWMSIYACRNASKAL 1052
Qy      1049 HVDLLHSILRSPMFPERTPSGNLVNFRSKELDTVDSDIPEVIMFGSLFENVIGACIVT 1108
Db      1053 HGQLTNILRAPMRFPTTTPGRIVNRFSGDISIVDDLLPQTLLSMMWCFGIGLTVMI 1112
Qy      1109 LLATPILAIIPPLGLIYFVQRFYVASRQLKLESVSRSPVSHNETLLGVSVIRAF 1168
Db      1113 CMATPVPALIIIPLSIYISVQVYVATSRQLRDLDSVTSKSPISHSEVETGPIIRAF 1172
Qy      1169 EEOGRFTHQSDLKVDENQKAYPISIVANRMLAVNLECVGNCIVLPALLPAVISHSISAG 1228
Db      1173 EHOORFLAMNEKQIDINQKCFSWITSNRMLAIRLELVGNLVVFCASLLVIYAKTILGCD 1232
Qy      1229 LVGLSVSYLQVTTYLMLVMSSEMETNIVAVERLKEYSETEKEAQMIOETAPSPSWP 1288
Db      1233 VGVFVLSNALNITOTLMLVMTSEATNIVAVRISBYINVEEAEPV-VTDKRPDPDWP 1291
Qy      1289 QVGEVEFRNYCLARYREDLDVFLRHINVTINGEKVGIIVRTGACKSSLTGLFPIINESAE 1348
Db      1292 RHGELIQNNYQVRRPELDVLKGITCNIKSGEKVGVGRGAGKSSLTNCLFRTLSAG 1351
Qy      1349 GEIITDINIAKIGLHLRFKTIITIPDDVLFSGSLAMNLDPSQSGDEEVTWLSLELAHL 1408
Db      1352 GQIITDIDIVASIGLHLRELTITIPDDPLFSGSLAMNLDPSQSGDEEVTWLSLELAHL 1411
Qy      1409 KDFVSAIPDKLDHCEAGGENTLVGQROLVCLARALRKTILVLDATAVADLETDDL 1468

```

```

Db      1412 RSFVSGQLGLSEVTEGDNLSIGQROLCLGRAVLRRKSLIVLDEATAVADLETDSL 1471
Qy      1469 OSTIRPOEDCTVLTIRHRLNTIMDTRVYVLVDKGEIOEGAPSDLIQOGLFYSMAKDA 1528
Db      1472 QTTIRKESQCTVITTAHRLHTIMDSKIMVLDNGKIVEGSPBELLSNRSFYLMKEA 1531
Qy      1529 GL 1530
Db      1532 GI 1533

RESULT 5
E89447
protein F57C12.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: E89447
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.el
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an
A:Accession: E89447
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1494 <STO>
A:Cross-references: UNIPROT:Q20943; UNIPARC:UPI000017801E; GB:chr_X; PIDN:AAA83299.1; P
A:Genetics:
A:Gene: F57C12.4
A:Map position: X
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match      32.8%; Score 3289; DB 2; Length 1494;
Best Local Similarity 47.6%; Pred. No. 1.2e-213;
Matches 696; Conservative 239; Mismatches 435; Indels 92; Gaps 18;

Qy      119 LAFILQIERKGVQSSGIMLTFWVALVYCALAIRSKIMTALKEDAQVDLFRDITRVY 178
Db      76 LAILTVACNKKGIITSGVITLVLVVGVGIEFRFYLGGFIYNEVALGIRATLVIIA 135
Qy      179 FSLILQIVLSCFSRDRPLSETIHDPNCPRESSASFLSITTFWVIGLIVRGYRPLEG 238
Db      136 FTSSALEFLCCPAD--VSDMYKSSSCPEYTAISTNTTQWFTGLVYLGKKSLEN 192
Qy      239 SDLMSLNKEDTSQVVPVIVKMKKECAKTRKQPVKVVSXDP--AQPKSSKVDANEB 296
Db      193 EDLMDNLEIDKAEMLDPSFQNLKPRIDEVH-QNIK-----KQPSALPKN----- 237
Qy      297 VEALIVKSPQKKNPISLFKVLVYTFGEFPLMSFFPRAIHDLMFSGQIILKLIKPVNDT 356
Db      238 -----HPSFVPIPFYKYKTYTLGAGFFYKLCFDMLOFLAPQLKOLIGFIEDK 284
Qy      357 KAPDMQGYFTVLLFVACIQTLVHQYPHICVSGRITVAVIGAYRKAIVTNSARK 416
Db      285 NQVWIGCSIVGIMFSSFLQSNFLHOYHSMRLGHAHVSVLTSAVSRALMINSERAK 344
Qy      417 SSTVGEIVNMSVDAQRFMDLATYINMWSAPLOVILALYMLNGSVLACVAVVLM 476
Db      345 GKTIQAIIVMSVDIQKIQDMPATIMFWGAPQIPLSIYFLWKFLGVALAGIVLILA 404
Qy      477 VPVNAVMAMKTKTYOVAMHKSQKNRIKLMNEIINGIKYLUKUYAMELAFKDKVLAIROEL 536
Db      405 LPVNGGLAIQMRKQTEOMKLKDERIKMSEIINGMKVLLKLYSMERSMENNVKIREREL 464
Qy      537 KYLKKSAVLSAVGFTWVCTPPLVALCTPAVYVITD-BNNIILAQRTAFVSLAFNIRFP 595
Db      465 HIILKLSYFMAALVFSICAPFLASVTSFVYVYIYLDENNVLPEITFVALSFDIIRMP 524
Qy      596 LNILPMVTSIYQASVSLKRLRIFLSHELEPDSIERRPYKDGCGTNSITVRNATPTWAR 655
Db      525 LAMVAVYGAVQCSVSNTRLKEFPAAEWSPOT---SISHGETDSAIRVENGLFWSMS 580

```

Qy	565	SDPPLNCTITPISIPBGALVAVVGVGCCSSILSLALAMDKYEGVNAIKGSAAVYPPQA	715
Db	581	DEDPFLREISFETIQGQVLAIVGVKSGKSSLHLALGEMNKLSGVQINGENIAVYPPQA	640
Qy	716	W1QNDLSRENILFFGCOLSEPPYRSYVQAQALPLDEILPSPGRTEIGEGVNISSGQOKR	775
Db	641	W1QNNLSLNNNILFPNRPYDENEDVYKNCALMEDLANLPAGRDTEIGEGKINLSGQOKR	700
Qy	776	VSLARAVYSNADIVYFPDDPLSAVDAAHVGHIFENVIQPK-GMLKMKTRILIVTHSMGYLPQ	834
Db	701	VSLAAVAVQNPDIILDDPLSAVDSHVGHIFENVISSSTGCCLASTKTRVLVTHGLTYLKH	760
Qy	835	VDVIVVMSGGKISEMGSVOELLARQDAFAEPLRTASTQEOBQAEENGTVGVSGPG----	890
Db	761	CDQLVLVBEGTISELGTYOELLNNSGAFAPLEELPESKTRGRVASYIGDGSGEVDELL	820
Qy	891	KEAKOMENGMLVTDGAKQOROLSSSS-----SVSGDISRH---NHSTBELQK	936
Db	821	RDLOGVKRGIL-----KRLSEHLGSESDKEPDSARAIETSRNRSRSVLIHSPRQOHE	874
Qy	937	AEA-----KKEETWKLMEADKAQOGVKLSVYWDYKAIQLPISLISIFLPMCNHV	987
Db	875	NEALIGALISEDVPAQENTOLIEKETVETGKVEKVEVIAVFOALISIDITLLFFFLVYGS	934
Qy	988	SALAASNNYMLMTDD-PIVNGTOEHKVMALSVYGALGISQGLAVPGCSMAVSGIGLISR	104
Db	935	LGILSNFYAKISDHAKSGNRTSSDAKMELGIAVAVGMGSPFVLLIASITLLTGVLPA	994
Qy	1047	CLHVDLLSHLILSPMSFFERTPESGMLVNRFSKELDPVDSMIPEVIKMFMSLFPNVIGACI	110
Db	995	ILHAGLLGNIMSPMAFPVTPPIGRILNNIGDIEAIDRLPVIINHMSTINNVAVTV	105
Qy	1107	VILLATPIAAIIIPPLGLIYFPVQRPYVASSROLKRLSEVSRSVPVYSHENETLLGVSVIR	116
Db	1055	VYMAATPMAIGIFALISVIFYFLRPIYISTSRQKRLSEASRSPIVSHFQESIQGASSIR	111
Qy	1167	APEEBERFHQSDKVDENOKAYYBETIVANRWMLAVLECGNCTVLEAALFAYISHS--	122
Db	1115	AFGVVDNFIKQSOQORVDHLIAIYPSIVANRWMLAVLEMGNTLVLSAQAAYFPDPSG	117
Qy	1225	LSAGLVGVSYSXLQVTTYLVNMLVRRSSSEMETNIVAVERKESYETKEAPMOIQETAPP	128
Db	1175	LSAGLVGLSVSALNITQOTLMAVNRKTSLEETNIVSVERKETYVPTGEG--NNSRRLAA	123
Qy	1285	SSWPOVGRVEFPNRYCLTRYEDDVLRIHNVITNGEKVGI VGRTAGKSSLTJLGRIN	134
Db	1233	KSMPEKGISIKNFSVRYRPGDLVYHGISAHIAPEKVKGIVGRTAGKSSLTJALFRIT	129
Qy	1345	ESABEGIIITDGINIAKIGHDLRFKTTIIPQDVLFGSSLRNMLDPDSQSDSEEWTSLE	140
Db	1293	EADGSGIEIDGINIANIAMEQLSCLTIIYQDVLFSGTMKMMIDPFSASDSQWEALE	135
Qy	1405	LAHLKDFVSLPDKLDHECAEGEENL-----SVGOROLVCLARALLRKTKI	145
Db	1353	NHHLKPFVYSLODGLSEHKISBEGGENLRFPAGSKQKQOLNFSVQOROLICARALLRKTIV	141
Qy	1451	LVLDENATAVDELTDLDLIGSTIRTOGEDCTVLTIAHRLNLTIMDYTRVIVLDDKEIOEYGA	151
Db	1413	LVLDENAAAADVETBDSLIQKTIREQPECTVLTAHRLNLTVMDSRDLVLVDKGRVAFSDS	147
Qy	1511	PBDLL-QORGLATYSMAKDAGLV	151
Db	1473	PNULANPDGIFYSMAKDAANV	1494

A:Reference number: Z19392
A:Accession: T21219
A>Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-1573 <W12>
A:Cross-references: UNIPROT:Q93552; UNIPARC:UPI00000823F7; EMBL:Z81016; PIDD:CA802667.1
A:Experimental source: clone F21G4
R:Kershaw, J.
submitted to the EMBL Data Library, March 1995
A:Reference number: Z19828
A:Accession: T24002
A>Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-1573 <W12>
A:Cross-references: UNIPARC:UPI00000823F7; EMBL:Z48621; PIDD:CAA88549.1; GSPDB:GND00028;
A:Experimental source: clone R07B1
C:Genetics:
A:Gene: CESP:F21G4.2
A:Map position: X
A:Intons: 21/3; 122/3; 197/3; 276/1; 654/1; 911/1; 1067/1; 1472/2; 1537/2
C:Superfamily: human multidrug resistance protein CMOA2; ATP-binding cassette homology

RESULT 6
T12119
hypothetical protein F21G4.2 - *Caenorhabditis elegans*
C.Species: *Caenorhabditis elegans*
C.Date: 15-Oct-1999 #sequence_rev: 15-Oct-1999 #next_change 09-Jul-2004
C.Accession: T12119; T24002
R.Mortimore, B.
submitted to the EMBL Data Library, October 1996

Query Match	Similarity	32.2%	Score 3225	DB 2:	Length 1573
Best Local	Similarity	42.4%	Pred. No.2.8e-209		
Matches	685	Conservative	305	Mismatches	490
				Indels	134
				Gaps	26
Qy	3	LRG---FCSADSDSDMDMDMNVMTWNSN--PDKFCFQONTLVWVPCFYLMACCFPFYLYLS	58		
Db	9	LKGIADAF---GGRYVD--PAVMNASTVPILISQCYOHTLVWVF-----	47		
Qy	59	RHDRCYIOMTPINKTKTVALGFLMIIVCWADLFYSF-----WER-----SR	98		
Db	48	-----TAIVFLIAPILTAQIFRYRRPNPIPMWRRIQLKIGLACILYADSL	91		
Qy	99	GIELAPVF-----LYSPULLGITTLTARFLQLERKQVQSSGIMLTFTLVVL	146		
Db	92	SLFTVAIYETTFQGFPPYADVFPYPLTLCLAMVLTALIVSCNNGVIVSGSLPISMLVPT	151		
Qy	147	VCALAIILR---SKINTALKEDAQVDFRDIYFYVFSLLIQLVLSCFSDRSPL--FSEYI	202		
Db	152	ISAIIBELLYWIOQIYNPAEMNMWIDYPRCIAFIIMFPCCAFETYLHCYADNAPBEQKYLS	211		
Qy	203	HDNPFCEPBSSASFSLRITFWITGLIVNGYROPLESGSDIMSLNKEDTSEQVYPVLVKNK	262		
Db	212	AARNSPSETTSSFLNRIITWMMFNNSLCISLGKKPFLRESDLYSINADTSLIVPKWYNLMD	271		
Qy	263	KECATRKQOPVNVYVSSSKDPQAPKSSKVDNAEEVEAL-----LYKSPQKEMN	310		
Db	272	KOSKKEETEAARRRIGSN--ASRTNRRRTSSDDTTPPLNDQSTDDGYSPACQSTQK--M	327		
Qy	311	PSLFLKLYTFCGPYFLMSFPFAIHDLMMFSGPOLIKLIRKVNPTKAPDMQGYFYTVLL	370		
Db	328	PSIITLFLFMFKWDVITAMFVLLSDVLLFCNPLIKSLIRTEELERFMMQGVLAFLFM	387		
Qy	371	FVTACIQTLVLIHQYFHCIVSGMRIKTAVIGAIVRKALVITNSARKSSIVGEIIVIMSDV	430		
Db	388	PSAISLSSILSHFYFLMYRGTNRVQTCITAAVYRKTLTLSNAREKTVGEIIVNLMAID	447		
Qy	431	AORFMDLATYIMIMSAPLOVILAYLMLNIGPSVLGAVAVYMLVPVPAVNAAMTKTY	490		
Db	448	VDRFOOITPQTOQWYNSNPFQISGIAELFLFQOQGVSVFSGVAWVLLFPINFIYTMIRKM	507		
Qy	491	OYAHMKSCKNRITKLMNEILNGIKVLKYVAMELAFKDKVLAIROBELKVLKKSAYLSAVGT	550		
Db	508	QIAQMYKQDERIKRMNEVINGIKVILKYIAMEPPEMEQVIDDLRQELGILKKAALFLTFSD	567		
Qy	551	FTWVCPPPLVALCTAFAVYVITIDENNILDAQTAFVSLATFNILIRFPLNIPMVISIVQAS	610		
Db	566	MLNTASPFVLVALSTFATFIYIDPKQVLPREIAFVSLTFCNQRSPMSQVAILITQVQVY	627		
Qy	611	VSILKRLIRLISHELEFPDISERRPVKDGCGTNSIYKRAFTFW--ARSDP--PLNKGITTS	667		
Db	628	VNRRIRKEVLVEEELINVEALIDHR--ARD--NNQVVICKEKCLSWESAEHQPVPPLTINIS	684		

Ckeywords: A1; nucleotide binding; P-loop; transmembrane protein; yeast vacuol

F:287-308/Domain: transmembrane #status predicted <TM1>
F:345-366/Domain: transmembrane #status predicted <TM2>
F:441-442/Domain: transmembrane #status predicted <TM3>
F:446-467/Domain: transmembrane #status predicted <TM4>
F:534-555/Domain: transmembrane #status predicted <TM5>
F:558-580/Domain: transmembrane #status predicted <TM6>
F:646-828/Domain: ATP-binding cassette homology <ABC1>
F:663-670/Region: nucleotide-binding motif A (P-loop)
F:951-972/Domain: transmembrane #status predicted <TM7>
F:995-1016/Domain: transmembrane #status predicted <TM8>
F:1068-1088/Domain: transmembrane #status predicted <TM9>
F:1092-1113/Domain: transmembrane #status predicted <TM10>
F:1119-1200/Domain: transmembrane #status predicted <TM11>
F:1208-1229/Domain: transmembrane #status predicted <TM12>
F:1289-1483/Domain: ATP-binding cassette homology <ABC2>
F:1306-1313/Region: nucleotide-binding motif A (P-loop)

Query Match 28.3%; Score 2838; DR 1; Length 1515;
Best Local Similarity 40.3%; Pred. No. 4e-183;
Matches 632; Conservative 295; Mismatches 495; Indels 146; Gaps 27;

D DFTKCFQNTVWVWCPCFLMACPPRY-----LYSRHD-----RGYIQMTPLTKTY 75
D 27 DFGCFIDGVL-----NLSAIFMTTFGRDLVNLCCKHSSIKRRMWIIYSRALVLL 81

O 76 ALGLFWIVCNADLFYEFWMERSRGIFLAPVFLVSPTLLGITTLATFLIQLERRKGQSS 135
D 82 ELAF----VSLASLNISKEAE--FPIVSGYSTMLSLPVALLAHWIEYDR--SVAN 132

O 136 GIMLTFWLVALVCALAIRSKIMTALKDAQVDLEFDITIFYFFS-----LLLIQLVLSG 190
D 133 TVLLFYWVFETFGNPAKLINILIRHTYG-----IWMSGQTGFILLTPGVITC 180

O 191 FS-----DRSP-----FSETHDPNPCCPSASPISRITPEMWTIGLVGYRPLEG 238
D 181 ASILLIELPKPFMPHOHIHOTLRKRKPNDSDANSITSRTIFSWSGMIAKTGEKYIVE 240

O 239 SDLMSLNKEDISEQVPVLVKNWKCEAKTRKOPKVUYSSKDPAQPKESKVDANEVE 298
D 241 ADVLYLPNFSEBELSQLEKNWENEL-----KQS----- 271

O 299 ALIVKSPQKENWPSPFKVLKTCGGYFLMSAFFKAIDLAMS GGQILKLIKFPVD--- 355
D 272 -----NPSLSMAICRTFGSKMLLAFAFKAIHVLAFTQPOLRLILKFYDVNS 320

O 356 -----TKAPDMOGYFTTVLLFNVAACLOTVLHOOYFHICFPGMKRIKA 398
D 321 ERDDHSLSQGFEHHNQPLPLVRGPLLIAFMFLVGTTSTVSHQFLPNVPMGTMIKSA 380

O 399 VLGAVYRKALVTITNGARKSSTVGEIVINIMSVDARQRMDATVINNIWASAPLOVIALLYL 458
D 381 LTAIIYQGSIVLSNASGSIISTGDIVNLSVDVQGLDTQWNLIMWGPRPDIICTLSEL 440

O 459 WLTIGPVLAGVAVNWLMVPUNAAMKTKITYOVAHMKS KDNRIKLMEINGIKVKLY 518
D 441 YGLIGNSMWVGVIILIVIMPINSFLMRIOKLQKQOMKXDEBTRVISIILNNIKSLKY 500

O 519 AMELAFCUKVLAIR-QEELIKYAKKASYLSA VGTFTFWCVCPREALCTFPAVVYTIDBNLL 577
D 501 AMEKYEKKRELBENVNKELKNTLKCGCYAAVTSFDFNVIPFLVSCCTFAVFA-YIEDRAI 559

O 578 DAQTAFFSLALEFINLRPFLINDPWYISSIVQA SVSLKRLRIFELSHHELEPPDSIERPPVD 637
D 560 TTDLVFPALTIFNLSSFPLMITPWLNSFI EASVSIGRLFTEFTTNBELOPDSVQRLPKX 619

O 638 GGGTINSITV-RNAETFWRSD--PPTLINGITESEGALVAVAVOVGCSSKSLIALAB 694
D 620 NIGDAVINIGDATEFLMRKPEYKALKNINQAKKGNLTICIVGKVGSGKTALLSCMGLD 679

O 635 MDKYEGHVAIKCSVAVYPQAOAITONDSLRENILFGCOLEEPYYRSYIOCALLPLEELP 754
D 680 IFRVVGFAIVHOSVAVYSGVPMWIMGVINEKITLPHGRDYDAEEYKTIKCAITIIDAILM 739


```

QY 1227 GGLVSLSYSSVQVTTNMTLVWMSSEMETNVAVERLKEYSETEKEA PMOIQETAPRSS 1286
DB 1173 SGLVSLSSVAAQVITQSLEFVVRQSDVDETNTVSERMLEYITGLPSESPSII PDHRRPEG 1232
QY 1287 WPOVVERFRRNCLCYREDLDFVLRIHINTVINGSEKGVIGRTGAKSALTGLPRISES 1346
DB 1233 WPSHAIRFEDHSAVYRRENLPLVINDISVNIKPQKIKIVGRTGAKSTLTALFRLIEP 1292
QY 1347 AEGEIIIDGINIAIKGLHDLRPKTIIPQDPVLFSGSLRMNLDPEFSQYSDERWTSLELA 1406
DB 1293 TSGDIQLDINDINTSIGLDHLSRSLAIIPQENQAFGRTIRENLDPMANNTDEIWHALAEA 1352
QY 1407 HLKDVSLPDLTDEHCAGEGENLSVGGOROLVCLARALKRTKIIVDEATPAVDLETDD 1466
DB 1353 SLKQPIQTLDDGLYGRVTEGGANLSSGGOROLMCLTRALLTPRVLIDDEATAVDVETDA 1412
QY 1467 LIQSTIRPFQEDCYVLTAAHRLNTMDYTRVVLIDKSHIOYGAASDILQOR-GLFYEMA 1525
DB 1413 IVQRTIRERENDRTILTAAHRLNTVMDSNRIIVLDHGKVEPDSYKLLKNKASLFYSLSA 1472
QY 1526 KDAGLV 1531
DB 1473 KESGLI 1478

RESULT 11
101369
ABC transporter AtWRP2 [imported] - Arabidopsis thaliana
N.Alternate names: multidrug-resistance protein homolog T29F13.13
C.Species: Arabidopsis thaliana (mouse-ear cress)
C.Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
A.Accession: T01369; D84759
R.Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, May 1998
A.Description: Arabidopsis thaliana chromosome II BAC T29F13 genomic sequence.
A.Reference number: Z14179
A.Accession: T01369
A.Status: translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 1-1623 <R0U>
A.Cross-references: UNIPROT:O64590; UNIPARC:UPI000009F2D5; EMBL:AC003096; NID:G3132469;
A.Experimental source: cultivar Columbia
R.Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Unayam, L.; Tallon,
euser, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-766, 1999
A.Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A.Reference number: AB4420; MUID:20083487; PMID:10617197
A.Accession: D84759
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-1623 <STO>
A.Cross-references: UNIPARC:UPI000009F2D5; GB:AE002093; NID:G3132479; PID:AACT6268.1; G
A.Genetics:
A:Gene: At2G34660; T29F13.13
A:Map position: 2
A:Introns: 110/3; 170/2; 183/3; 236/1; 274/2; 301/2; 309/3; 328/3; 350/3; 375/3; 411/3;
F:1259-1452/Domain: ATP-binding cassette homology ABCA2
Query Match 23.8%; Score 2384.5; DB 2; Length 1623;
Beat Local Similarity 35.2%; Pred. No. 2-2e-15;
Matches 561; Conservative 274; Mismatches 534; Indels 223; Gaps 30;
QY 44 FYLWACPFY-VELYLSRRDRGYIQMTPLNKTALG--FLMTVCNADLFYSFW----- 94
DB 5 FIEWCKEVPNGVNTKQYANAFGAYTPCATDSFVLGISQVLVILVC-----LYRIWLAKD 60
QY 95 -----EKSRLGFLA-----PVLVSPFLILGTTL-----LATFLIQLERRK 130
DB 61 HKVERFCLRSRLVNYFLLALAAVATAEPLFL-:::IMGISVLDDGPGCLPFP----- 109
QY 131 GVQSSGIMLTFWLVALVCALAILRSKIMTALKEDAOVDLPFDIFFYVFS-----LL 182

```

Db	110	--EAGLGVKAFAMGAVVMVIMMETYK-----YIRLRYVRRAVYALVGDWV	156
Qy	183	LIQVLVS-----CFSDRSPLESETIHD----	204
Db	157	LLNLVLSTKVEYSSVVLVLYTSEVGAQVLFGLLPMHLPNLDITYGCMVPRSETJDYVY	216
Qy	205	-----PNPCBSSASPLSRITTFMWITGLVNGRYOPLEGSDLWLNKEDTSEQVVPVLV	258
Db	217	BEISDGOQICECEKHANIFDKIFPFSWMNPLMTGSRPLTEKDVWYLDPTDQETELFTSPQ	276
Qy	259	KNWKECAKTKRQPKVYVYSSKDPAPQPKSSRYVDANEVEALIVSPQKEMNPSPFKVLY	318
Db	277	HSWDELQK-----PQ-----PWLIRALN	295
Qy	319	KTFGPFYFMSFEFFKAIHDLMFSGPOLKLLIKFVNDKRAPDMOGYFTVLLFPAACLOT	378
Db	236	NSLGRFPMWGFWKJGNDCSQPVGPBLLNOLIKSQOE-DAPMMGYIYAFSLFVGVRFEV	354
Qy	379	LVLDYFPHICEVSGNRIKTAIVAGYVRKALVYTNARSKSSTGEIVNLMSYDAQRFMIDA	438
Db	355	LCEAQYFQNMWRVGRRLRSALIAAVFRKSRLTNEGRKRFQGGKITNLMTTDAESLQOC	414
Qy	439	TYINMWSAPRLQVILALYLLMLNIGPSTIAGVAVWVLTVPVNAVMAKTKTYQVAMHKS	498
Db	415	QSLHTMWSAPRRIIATLILYQOGLVNSALIGALLVIMFPLQTVIISKOQKLTKEGLQRT	474
Qy	499	DNRILKMEIINGIKVLKYAMELAFKQVLAIRQELKVLKKSAYLSAVGTFTWCPRF	558
Db	475	DKRIGLMEVLAANDTVKCIAMENSFOKQVOTRDELSWFRKSQULGALNMFIANSIDV	534
Qy	559	LVALCTPAVYTYTIDENNILDAQTAFFVSJALFNILRFPNILLPMVSIISIVQASVLSKRL--	616
Db	535	LVTIVSBEVFTLGGD--LTPARAFSTLSLFAVLRFPFLPMLPNITQYVYNAVNSLKRLEE	592
Qy	617	-----RIFLSHELEPDSIERRPYKQGGTNSITVRNATFTW--ARSOPPTINGITFSIP	669
Db	593	VLAITEERILLENPIEP-----GEPALISIRNGYSWMSKGRPFLTSLNINLDP	640
Qy	670	EGALVAVVGVQCGCKSSLLSALLLAEMDKV--BGHVAIKGSVAVVYPOQAINONDSLBNILF	728
Db	641	LGSVLAVVVGSTGEGKTLISALIGLEPATSPAIYTLKGSVAVVPOVSVI FNAVTRDNILF	700
Qy	729	GCQLEBPYRSVIOACALLPDLLEILPESGDRTEIGEKVNLGGQQRORSALARAAYSNADI	788
Db	701	GSPDRKREKERAIDVTSLKHDLELLPGSDLREIGRGVNSIGGQQRORSAMARAYVSNBDV	760
Qy	789	YLPDDPLSAVDAAHYGKAI FENVYIGBKGMKKNKTRILVTHSMSYLPQVDVILVMSGKISE	848
Db	761	YIFDDPLSALDAHQOQVFKECI--KELGQKTRVLVYNQHLFLFSQVDRIVLVHEGTAYKE	818
Qy	849	MGSYQELLARCGARAEPLRTAYTSBQEDBAENGVTVGSGQKE--AKOMENGMLVTBQA	906
Db	819	EGTYEELSSNGPLRQRLMENAGKVEEY--SEENGEAEADQTAEOPVNANGNTNGLOMDSGD	876
Qy	907	GKQYORQOLSSSSYSYSDISRHHNSTABLOKAEAKKEETWKLMKADKAOTGOVKSIVWYD	966
Db	877	DKK-----SKGNKKGGKSULIKQEBERTGVVSNRVLKRY	911
Qy	967	MKAI--GLFISPLSIFLPMCNHVSAALASNYWLSLWTDPEIVNGT--QEHTRKVLRS--VYGALG	1023
Db	912	QDALGGAVMVMMLLICVLTLEVFVTSJTWLSEWTD----AGTPKSHGRLPFLYNLITALLS	967
Qy	1024	ISQCIANFGYMAVSIIGIILASRCLHVDLHSLIIRSPMSFEFERPDSGLVNRFSKEDTV	1083
Db	968	FGQVLAVTTLTNSYTWIMSSLYAKKLHDMHLSIRAPMSFHTNPLGRI INRFADLDGI	1023
Qy	1084	DSMIPVILKMGMSLFWNIGACIYIILATPIAAILIPPLGIFTFVFGVYAASSROLKRL	1144
Db	1028	DRTAVAVFMNMGQVQSLLSTVVLIGIVSTLSWAINPDLVLFPYGALYYIYQNTAREVGRM	1083
Qy	1144	ESVRSRPVYSHFNELLIGSVITRAPEOEERFIHOSDLKVDENQKAYPSIVANMLAVRL	1203

Db	1088	DSISRSPVYAAFGELANLSTIRAKYADNRMAIDNGRBMDDNNIRPTLVYMGANRMLGRRL	1147
Qy	1204	ECVANGCIVLFALPAVTL-----SRHSLSAGLVGSVSYSLQVTTYLANLVKMSSEMETNI	1258
Db	1148	ETLGGIMLWMLFASFPVAMONGRAENQOAFASITMGILLSYALNITISLTITGLVLRSLAESNSL	1207
Qy	1259	VAVERLKEVSETEKRAPWQIQIETAPSSMPQVGRPEFNQVCLARXEDDLPFLRHINVTIN	1318
Db	1208	NAVERGVNYITIPRAPPVYIENNRPPGMPWSSGSIKFEEDVVLRYRQPLPPVLHGVSEPTI	1267
Qy	1319	GGEKVGIYGRTAGCKSSLTLLGFRINESAGEGIIIDGINIAKIGLHDLRFKITIITPOPV	1378
Db	1268	PTDKVGIYGRGACKSSLLNMLFRIVEVEKGRITLDDCCDVGKFGMLDRKVIGIIPQSPV	1327
Qy	1379	LFSGSLRNMULDPSQSYDDEBVTSLLELAHLKDPFVALPDKLDHCEACGEGENLVSQROLV	1438
Db	1328	LFSGIVTRNULDPSGHNADALMESLERAHMLKDTIRRNLLGIDABVSEANGENPFSQGRDLL	1387
Qy	1439	CLARALLRKTKILVLDEATPAVDELTDDLQSTIRTFQEDCTVLTIARLNTIMDYTRVI	1498
Db	1388	SLSLALLRSTRKILVLDEATPAVDVTDLLIOKTIRFEFKSCMTLLIARLNTIIDCKIL	1447
Qy	1499	VLDKGEIOEYCAPSDLLQORGLFVS-MAKDAG	1529
Db	1448	VLDGSRVQVDFSSPENLLISNEGSSFSKMWQSTG	1479

```

OY      224  ITGIVNGQPLBEGSLMLNKEDTSEQVVPVLVNMKKCECATRQOPVXVYSSDPA 283
Db      242  LNPIMTUGSKRPLTEKQVMLDWTDKETLMRSFQSKWDXLEK-----285
OY      284  QPKESKVDANEEVEALIVKSPQKEMNPSLFKVLYKFKGFPYILMSFFPKA1H1DLMFSGP 343
Db      286  -PK-----PWLRLANNLSGGRFPMWGGRWKIGNDSQGVGP 320
OY      344  QILKLLIKFVNDTPAPDMQGFYTVLLFVTAQIQTIVLHQYFHC1FVSGMR1KTAVIGAV 403
Db      321  LLNLNELKSM-Q1NEPAMIGYIYASIFGVGVGLVCEAQGFQVWMEVGR1RSAL1AAV 379
OY      404  YRKALVITNSARKSSYTGELIVNLSVDAPQRPMDATTINNIWSP1LOY1IALY1LMLNLG 463
Db      380  FRKSLIRLTNEGRKKFQGGKITMLNTTDAESIQQICOS1HTMWSAPFR1IYALV1LXYOOLG 439
OY      464  PSVLGAVVAVMLAMPVNAVMAMKTKTYOVAMHMSKNDR1KLMNS1LNG1KVLKYAMELA 523
Db      440  VAS11GALFVLMPPIQTIV11SKQKLTKEGLQPTDR1G1MNEV1LAMDTV1KCYAMENS 499
OY      524  FKDKYLA1ROBELVYLLKKSAYLSAVGTFTWVC1PFLVALCTFA1YVY1DEN1N1DAOTA1F 583
Db      500  FQSKVQYVTRDELSWPKRAQLLSAFNMFIINS1FVLT1VTVVSFGVFS1LGGD--LT1PARAF 557
OY      584  VSLAFEN1LRFPLN1LPMV1SS1VO4SVSLKRLR1F1LSHEE--LEPDS1ERRPVYDGGGT 641

```

RESULT 12
D864428
glutathione S-conjugate transporting ATPase (AtGMR1) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: D864428
R:Thelloglis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Hultzar, L.
Nature 408, 815-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzilli
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzbach, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D864428
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1622 <STO>
A:Cross-references: UNIPROT:Q9CGG9; UNIPARC:UPI000009CF47; GB:AE005172; NID:g11055814; F
C:Genetics:
A:Map position: 1
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match	23.5%	Score 2356.5	DB 2	Length 1622
Best Local Similarity	35.6%	Pred. No. 1.8e-150		
Matches	569	Conservative 274	Mismatches 564	Indels 193
			Gaps	31
QY	47	WACPF-YELVLSRDRGIQNTPLNKKTKALG-----FLMIVICADLFYSFW	94	
DB	8	WYCFEPVNGWMTKYVDYAFGAYTCAIDSPVIGISHLVLLICTRLMLITKHKKDKFC	67	
QY	95	ERSR-GIFLA-----PVFLVSPTLIGITTL-----LATFLIQERRKQVS	134	
DB	68	LRSKKFSYFLMLLAAYATAPLFRL---VMRISVLDLDGAGFPPEAFLMVLVE--AFAMG	122	
QY	135	SGIMLTF-----WLVALCALAILRSKIMTALKEDAQVDLPDITFYFYFSLL	183	
DB	123	SALVMTVETKTYIHLELRMYVFAVIYALVGDVMILNVLVSK--EYSGSFXYIYSIVA	181	
QY	184	IQLV-----LSCFSDRSPLFSEFIHD-----PNPPRESSASGLSITFWW	223	
DB	182	VQVAFGLLLFYVFPNLDPPGYTTPVGTENSESEYEEELPGGENICPERRHANLPDSIFESW	241	

[illegible]

Db 1232 PSSSIFKEDVAVLRYRELPEVLHGVSFLISPMDKVGVIGRTGAGKSLNALPRIYELE 1291
 Qy 1348 EGEIINDGINIAKGLHDLRKTITIPQDPVLFSGSLAMNIDPFSQYSDSEVMTSLBLAH 1407
 Db 1292 KGRLLIDECIDIGRGLMDLRKVLGIIPAPVLFSGYVAFNIDPPSEHNDALWESLEBAH 1351
 Qy 1408 LKDFVSALPDKLDECAEGENGLSVGQROLVCLARALLRKTIIIVLBETAADVLETDL 1467
 Db 1352 LKOTIRRNPLGDLDEVTAGENFSVGQROLSLARALLRKSIIIVLBETAADVLETDV 1411
 Qy 1468 IOSTIRQFEDCTVLTIAHRLNTIMDYTRVIVLDKGEIOEYGAESDILQGRGLFYSMAKD 1527
 Db 1412 IOKTIRREFSKCTMLIIAHRLNTIIDCCKVLVDGKVOEFSFSPENLISNESSFSK--- 1468
 Qy 1528 AGLVGGGGGMLSKRGKIIPEYVULTRLEADPAERIRYTR 1567
 Db 1469 -----MVOSTGTANMEYLRSLTLENK-----RTRE 1493

RESULT 13

E86428
 Probable ABC transporter [imported] - Arabidopsis thaliana
 C/Spectes: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C/Accession: E86428

R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 Ansen, N.F.; Hughes, B.; Hultzer, L.
 Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A/Reference number: A86141; MUID:21016719; PMID:11130712

A/Accession: E86428
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-1495 <STO>
 A/Cross-references: UNIPROT:Q9C8H0; UNIPARC:UPI00000482FF; GB:AE005172; NID:g11055818; F
 C/Genetics:
 A/Map position: 1
 C/Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 23.24; Score 2326; DB 2; Length 1495;
 Best Local Similarity 34.34; Pred. No. 1.8e-148;
 Matches 541; Conservative 314; Mismatches 566; Indels 154; Gaps 28;

Qy 2 ALRGFGS-ADGSDPLMDANVTNNTSNPDTKCPQNTVLVWVP-----CFY-LMACFP 51
 Db 5 ALMWYCPVADGF-----WEKAVDGAFCAGATPCADISLWLVSHFVLLGLCFYRWITIF- 58
 Qy 52 FYFLVLSRHDGXYIOWTPLKTKTALGFLIMYQWAD-----LFVSFEMRSRGIFLAPV 105
 Db 59 -----HNKKAQIYLRKKYINCVLG-LACYCVAEPVRLVWGLSFLPMDETDPPF 110
 Qy 106 FLVSPFLIGITLLAFLLQLERRKGVSGSIMLTFWLVALVCALAIRSKIMTALKEA 165
 Db 111 EVASLWAEAFWFSMLVLIGLETRKQYKEFRWYVRFVLVVAADVLLDLVLPKNSIN 170
 Qy 166 QVULFADIT---FYVFSLLLIQVLVSCFSRSPFLSETTHDP-----NCP 209
 Db 171 RYALVFISRRGQALFGILLIYIPEL--DYPFGYIVNNEPLDNEVDALRGGEIICP 228
 Qy 210 ESSASFLSRITFMWITGLIVRGYROPLEGSDLSINKEDTSEGVVPLVKNWKECAKTR 269
 Db 229 ERASISFRIYFPGMITPLMGLGRKPTTEKDWQDLKMDOTETLIKRFQGCWIBESR-- 286
 Qy 270 KQPVKVVYSKDPAPQKSSKVDANEVEALLVKSPOKEMNPFLFKVLYTTFGPYFLMSF 329
 Db 287 -----PK-----PMLLRALNNLSLGGRFWLAG 307

Qy 330 FFKAIDLMFMSGPQILKLIKPVNDIKAPDMQGYFTVLIFVTACLOTVLHGYFHICF 389
 Db 308 IFKIGNDLSQVGVVILSHLSRMOGED-PAWGYVAFLIFVGTVLGVCEADYFQNW 366
 Qy 390 VSGMRITAAVGAAYRRALVYTNARSARKSTVGEIVNLSYDAQRFMDLATYNNIWSAPL 449
 Db 367 RVGRRLASTVVAALFHSLSLTHEARKNPASGKVTMTITDANLQOISQHLMSARF 426
 Qy 450 QVIALYLLMLNIGPSVLAQVAVNVLVNVAVNMAKTKTYQVAMKSKDNRIKLMNEIL 509
 Db 427 RIIVSMILLYQQLGVASLFSGLIFILFIPIQTLISKRLTKRGLQWTDKRVGITNEIL 486
 Qy 510 NGIKVLYLAMELAKFKQVLAIRQELKVLKSKAYISAVGTFVTCRPLVALCTPRAVY 569
 Db 487 SSMDTVACVAMKESFESRIGINNEBELSWPRKQLSAFNPSFILNSIPVAVTVVSGFV 546
 Qy 570 TIDENNILDAQTAFAVSIALFNILRFPNLIIPMVSSIVQASVSLKRL-RIFLSHELEPD 628
 Db 547 ILGCD--LTPARAFTSLFAVLRFPLMLPNLSQVYVNAVSLQRLBELLSERILAQ 604
 Qy 629 SIERRPYKDGGSNTSVRNATFTW-ARSDPPLNGITFSIPGALVAVVQVCGKSL 687
 Db 605 NPLQF-----GTPAISIKNGYFSWDSKTKPTLSDINLEIPVGTVAIVGTEGKSTL 659
 Qy 688 LSLALLMEMDVE-CHVAIKSVAVVPOQAMIQNDLSRENILFGQLEPPTRYIOCAL 746
 Db 660 ISMLGLSHAETTSVVRISVAVVPOVSVIENATVARENILFGSDFESERYRAIDATAL 719
 Qy 747 LPLEILPSPDRTIEGKGVNLGGQKORVSLRAVYNSADIVLPDPLSAVDPAHVAKHI 806
 Db 720 QHIDILPGLDLIEIGRGVNISSGQKORVSMARAVYNSDVYIFDPLSALDPAHVAV 779
 Qy 807 FENVIGKGMKNKTRILVTHSMSTYLPQVDVVIIVMSGKISBMSYOEELARQAPAEFL 866
 Db 780 FDSOM-KDRLRGKTRVLVNLQHPFLPMKXIIIVSGMKEBGTFLVLSKSLFLFKLM 837
 Qy 867 RTYASTBQEDAEENGVYVSGPGKEAKQENGLVYDSAGKOLQROLSSSSSYSGDISR 926
 Db 838 ENAGKDATQEVNTN-----DENILKGPVTVVDS---ERNIGSTK----- 876
 Qy 927 HHNSTAELQAEAKKETWLTMEADRAQOTGVLSYWDYMKAI-GLFISFLIFLFCN 985
 Db 877 -----QGRKRS-VLIKQERETGISMNVLAKYKAVGGIMVMIILACTYAT 924
 Qy 986 HVSALASNWYLSMTDPIVNGTOEHTK-----VRLSVGALGISQIAVGYSMVSG 1040
 Db 925 EVLRVSSSTWLSITWD-----QSTSKVYSPGYIVVVALDFGQVAVFTMSFMLITS 977
 Qy 1041 GILASRCLHVDLHSLIRSPMSFFERTPSGNLVNRSKELDYDSMIPVYIKRPMGSLFN 1100
 Db 978 SLHAARLHDAMSLSLRAFMLEFHNTPTRVIRNRSKIDGIDRRVAVANIMNFMNQLMQ 1037
 Qy 1101 VIGACIVIIATPIAIIIPGLIYFPVQRFVAVASROKRLSEVSRSPVYSHENETL 1160
 Db 1038 LLSFTPLIGVSTISLMAIPLDLILFYAALVYQSTSRERKRLDSVRSYVQFGBALN 1097
 Qy 1161 GVSIVAFEEBERFIHQSLDKVDENOKAYPSIVARNMLAVRLCEVCNCTLPAALFAVI 1220
 Db 1098 GUSIRIAYKAYDMARKNGSMONNIRFTLANSSNRMLTIRLETIGGVWIMLTATPAVL 1157
 Qy 1221 SRHSLS-----AGVLGVSYSIQVTTYLVNMLVRMSSEMETNIVAVRLKEYSETEKEAP 1275
 Db 1158 QNGNTNQAQGFASITWGLISLTITNLSLGSGLVROASRAENSINSYVERVANYIDLSEAT 1217
 Qy 1276 WQIQETAPSSQVQVGRVFRNRYCLRYREDLVDVLHINVTINGKEVGVVGTGTGAGKS 1335
 Db 1218 DLENNRPVCGWSSGSIKFEVDHLKRPOLPVLVGLITFPVSPSEKVGVTGTGAGKS 1277
 Qy 1336 LYLGLFRINESAGEIIDGINIAKGLHDLRKTITIPQDPVLFSGSLAMNIDPFSQYS 1395
 Db 1278 MLNALPRIIVEKGRIMIDCDVAKFGLTDVRVLSIIPQSPVLFGTVAFNIDPSEHN 1337

A;Residues: 1-255 <BOY>
A;Cross-references: UNIPARC:UPI00001746CF
A;Note: the authors translated the codon CAG for residue 248 as His
R;Purnelle, B.; Goffeau, A.
submitted to the EMBL Data Library, April 1996
A;Description: The sequence of 32 kb on the left arm of yeast chromosome XII reveals 14
mily and a new ABC transporter homologous to the human multidrug resistance protein.
A;Reference number: S69380
A;Accession: S69391
A;Molecule type: DNA
A;Residues: 1-1559 <PDR>
A;Cross-references: UNIPARC:UPI0000126A9C; EMBL:X97560; NID:g1297003; PIDN:CAA62776.1; F
R;Miosga, T.; Zimmermann, F. K.
Yeast 12, 693-708, 1996
A;Title: Sequence analysis of the CEN12 region of *Saccharomyces cerevisiae* on a 43.7 kb
e conductance regulator protein CTR.
A;Reference number: S70557; MUID:96405918; PMID:8810043
A;Accession: S70560
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1559 <MIM>
A;Cross-references: UNIPARC:UPI0000126A9C; EMBL:X91488; NID:g1495203; PIDN:CAA62776.1; F
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1995
C;Genetics:
A;Gene: SGD:BPT1; MIPS:YLL015W
A;Cross-references: SGD:S0003938
A;Map position: 12L
C;Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology
C;Keywords: ATP; nucleotide binding; P-loop; transmembrane protein
F;28-44/Domain: transmembrane #status predicted <TM1>
F;143-159/Domain: transmembrane #status predicted <TM2>
F;178-194/Domain: transmembrane #status predicted <TM3>
F;334-350/Domain: transmembrane #status predicted <TM4>
F;421-437/Domain: transmembrane #status predicted <TM5>
F;526-542/Domain: transmembrane #status predicted <TM6>
F;550-566/Domain: transmembrane #status predicted <TM7>
F;654-647/Domain: ATP-binding cassette homology <ABC1>
F;672-679/Region: nucleotide-binding motif A (P-loop)
F;974-990/Domain: transmembrane #status predicted <TM8>
F;1017-1033/Domain: transmembrane #status predicted <TM9>
F;1099-1115/Domain: transmembrane #status predicted <TM10>
F;1118-1134/Domain: transmembrane #status predicted <TM11>
F;1212-1228/Domain: transmembrane #status predicted <TM12>
F;1319-1529/Domain: ATP-binding cassette homology <ABC2>
F;1336-1343/Region: nucleotide-binding motif A (P-loop)

Query Match 22.9%; Score 2294.5; DB 1; Length 1559;
Best Local Similarity 35.1%; Pred. No. 2.6e-146;
Matches 574; Conservative 292; Mismatches 541; Indels 227; Gaps 39;

QY 23 NTSPDPTKCPONTVAVWPCFYLM-ACFPFYLYLSRHDRGYIQMTPLNKTATLQFL- 80
DB 23 NALNP-----CFIVISAMQAVFLLIGSYQLMKLYKNNKVPFRPKNPTLPKINSRHLT 78
QY 81 -LWIVCW-----ADLFYFMRBSRGIFLAPVELSPULLGI---TTLIAT 121
DB 79 HLTNVCQSTLIICELALVQSSDRYPYTLK-KALYTLNLF-----NLGSLPTQYLA 132
QY 122 FLIQLERRKVOSSGIMLTFMLVALVCALILRSKI-----MTALKEDAOV 167
DB 133 F-----KSTFSGNGNQLFYMPQIILQFLILGRYHSSNERTLVISQGTAMILEVL 165
QY 168 DLFRDITFYVFSLLIQVLVSCSDRSPLEFSETHDPNCPRESSASFLSRITFWITGL 227
DB 186 -LNSVAIFLY-DLCIFEPI-----NELSEYKKNGWY-----PVHLSYITFIWNKL 233
QY 228 IYVGYRPLEGSDMLSLNK--EDTSQGVNPLYVNGMKKECAKTRKQGVKVVYSSKQDAP 265
DB 234 IVEIYR-----NKAIKDPNQLPPEVDNLK----- 259
QY 286 KESSKYDANEVEALIVKSPQKWNPSLKFVLYKTPFPYFLMSFFFAHIDLMMFSGPQI 345
DB 260 SISKEFANWELKMLNRN-----SLMRAIWKSFRTISVAMLYETSDLSVQPOF 312

QY 346 IKLIIKEVN---DTKAPDMQGYFTYVLLFVTACQLTLLQYFHCIFVSGMRKTAIVGA 402
DB 313 LRITDGLNPTSKYKPLPLNGVFLATLFLVYSVSVLTNOFYGIEPAGIGIGISLASL 372
QY 403 VYRKALVITSARKSTVGEIVNLMSYDA---QRFMOLAATYINMISAPLOVILALYLLW 459
DB 373 VYQSLSLTLAERNKSTGDIINLMSVDVLRIGQFFENA---QTIIAGPIQIIVLISLY 429
QY 460 LNIQPSVLAVAVWVLPVNAVAMAKTKYQVAMHKSCKNRILKLNBIINGIVLYKLYA 519
DB 430 WLIGKAVIGGLVTMAIMMPINAFISRYKKLSKTQMKYKDMIRITIELNAIKSIXLYA 489
QY 520 MELAFKQVLAIRE-ELKYLKKSAYLSAVGTFTWCTPPLVALCTEPAVYVTTDENILD 578
DB 490 WEEMPMARLHVNRKDMELKNFRKGIYSNLIYFAMNCVPLMWTCTSTGLP-SLPSDPLS 548
QY 579 AQTAFVSLALFNILRPLNLIPLWYISSIVQASVSLKRLRFLSHELEPDSIER-RPVKD 637
DB 549 PAIVFPLSLFNILNSAIYVSPMINITIFTSVSMERLKSFLSDSIDSTIERIDISAD 608
QY 638 GGGNSTITVRNAFTW-----ARSDPPTLNGIT-----FSIPEGALV 674
DB 609 ERALPFAIEMNNITFLMKSKSVLTSSQSGDMILRTDEBSIIIGSSQIALKNIDHFEAKRGDLV 668
QY 675 AVYQVCCGKSSLSALLAEMDKYEGH-----VAIKSVAVVPOQAMTQNDSLAENT 726
DB 669 CVGRVAGKSTFLKALIGLPQWSGSRDSIIPKLLIRSSSVAACQSESMITMASVRENI 728
QY 727 LFGQLEPEPYRSYIQCALLPDLPIEPSGDREIGEGVNLSCGQKQRYSLARVYNSA 786
DB 729 LFGHKFPQDYDDLTIRKCOLLPDLKILPDGDELTVGSKGISLSCGQAKRSLARVYNSA 788
QY 787 DIYLFDDPLSAVDAVHGKHLFENV-IGPKMKLNKTRILVTHSNLYDPQVDVITVMSGK 845
DB 789 DIYLLDDIISAVDAVSGKNIIEYVLICKTALNKKITILLTNVYSILKHSQMIVALENGE 848
QY 846 ISEKSGYQELIAR---DGAFAELPRTYAS-----TEQDAEENGVTVSGPGKEAKOME 897
DB 849 IVQGNVEDVWNRKNNTSKLKLLEPDSIDNGNESDVOTERHSESEVEBP----- 900
QY 898 NGMLVTDASAKQOROLSSSSYSGDSIRHNSYAELOKAEKKEBTWKL----- 947
DB 901 -----LQKTBSETEDEVYT---ESBLIKANSRSLATLRPRPFVGAOL 945
QY 948 -----MEADKAQOTGVKLSYVWYMKAGLFIISLIFLM-CNHVSLASNYWLSLW 999
DB 946 DSVKTAQKAEKTEVGRVKTIVLAIYAKAGV-LGVVLPFLMILTRVFDLAEFMLKYW 1004
QY 1000 TDDPIYNGOEHTRKRLSVYGAIGISQGIAPFGYSMAVSI---GGLASRCLHYDLHSI 1056
DB 1005 SESENEKSGNERWMMFQVYSLIGVAS--AAFNNLRSIMMLVLCISIRGSKLHESMAKSV 1062
QY 1057 LRSPMSPFERTPSGNLVNRFSKELDPTVDSMIPEYIKMFGSLFNVGACIVILLAPPIA 1116
DB 1063 IRSPMTFFETTPVGRILINRSSPMDAVDSMLQITIFSEFFSILITVLTIVLVGYNMFWL 1122
QY 1117 IIPPLGLIYFYQRFVYVASSRQIKRLLESVRSRPSVSHFNETLIGVSVIRAFEOERFIH 1176
DB 1123 VFMNPLVLIYIYQTFYIVLSRELKRLISISYSPIMSLSMESLNGSIIIDAYDHPEFIY 1182
QY 1177 QSLKXVDENQKAYYPSIVANRMLAVRLECYGNCIYLPALPAVI---SRHSLSAGVLGS 1233
DB 1183 LNYEKKIQYVNDVFVFNFRSTNRKWSVLQITIGATIVATATIALATAMNTRKQLSSGVGL 1242
QY 1234 VSYSLQVTVLNMVLVMSSEMETNIVAVERTLKYSESEKAPQOETAPRSPQVGRV 1293
DB 1243 MYSISLEVYGLTIIVTTVTIETNIVSVEHIVYCELPPEAGSINBEKRPDEMPPSKGCI 1302
QY 1294 EFRNVCILRYEEDLPVLRHINVTINGEKKYIGRTGAGKSSJTLGLFRINESAEGEIII 1353
DB 1303 EFKNYSYTKRENDDPVLNNINIVKIBCEKQKIVGRIGAGKSTLSLALFRILEPTEGKIII 1362

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2005, 14:56:46 ; Search time 245.527 Seconds
(without alignments)
5594.767 Million cell updates/sec

Title: US-10-665-283-8
Perfect score: 10016
Sequence: 1 MALRGFCASADSPDPLMDMN.....RSVAVAKAKPKFSISPSLS 1947

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5.

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7860	78.5	1531	MRP1_HUMAN	P33527 homo sapien
2	7769	77.6	1515	09UQ39_HUMAN	Q9UQ39 homo sapien
3	7695	76.8	1531	0864R9_MACFA	Q864R9 macaca fasc
4	7689	76.8	1531	0864S0_MACFA	Q864S0 macaca fasc
5	7444	74.3	1459	09UQ97_HUMAN	Q9UQ97 homo sapien
6	7419.5	74.1	1456	09UQAO_HUMAN	Q9UQAO homo sapien
7	7272	72.6	1531	06UR05_CANFA	Q6UR05 canis famil
8	7242	72.3	1439	059G19_HUMAN	Q59G19 homo sapien
9	7230.5	72.2	1530	08HX05_BOVIN	Q8HX05 bos taurus
10	7104.5	70.9	1400	09UQ98_HUMAN	Q9UQ98 homo sapien
11	7002.5	69.9	1528	MRP1_MOUSE	Q35379 mus musculu
12	6932.5	69.2	1532	0810B4_RAT	Q810B4 rattus norv
13	6932.5	69.2	1532	08CG09_RAT	Q8CG09 rattus norv
14	6892	68.8	1523	0810G9_RAT	Q810G9 rattus norv
15	6301	62.9	1523	05F364_CHICK	Q5F364 gallus gall
16	6075	60.7	1215	068CP7_HUMAN	Q68CP7 homo sapien
17	4526	45.2	1522	05SUF4_MOUSE	Q5SUF4 mus musculu
18	4522	45.1	1519	080ZK8_MOUSE	Q80ZK8 mus musculu
19	4522	45.1	1522	056PH0_MOUSE	Q56PH0 mus musculu
20	4508.5	45.0	1523	059DU0_MOUSE	Q59DU0 mus musculu
21	4481.5	44.8	1527	MRP3_HUMAN	Q15433 homo sapien
22	4481.5	44.8	1533	059H05_HUMAN	Q59H05 homo sapien
23	4422	44.1	1514	096Q09_HUMAN	Q96Q09 homo sapien
24	4402	43.9	1498	059DK9_MOUSE	Q59DK9 mus musculu
25	4396	43.9	1522	MRP3_RAT	Q4S7B7 rattus norv
26	4148.5	41.4	1496	04S7B7_TRTNG	Q4S7B7 tetraodon n
27	3885.5	38.8	1548	Q7KTC3_DROME	Q7KTC3 drosophila
28	3880.5	38.7	1548	Q7KTC1_DROME	Q7KTC1 drosophila
29	3879.5	38.7	1548	Q7KTC2_DROME	Q7KTC2 drosophila
30	3874.5	38.7	1548	Q7KTC2_DROME	Q7KTC2 drosophila
31	3860.5	38.5	1548	Q917NO_DROME	Q917NO drosophila

32	3857	38.5	1549	2	Q7KTC0_DROME	Q7KTC0 drosophila
33	3845.5	38.4	1548	2	08T9C5_DROME	08T9C5 drosophila
34	3838.5	38.3	1548	2	Q7KTB8_DROME	Q7KTB8 drosophila
35	3836.5	38.3	1548	2	Q7KTB9_DROME	Q7KTB9 drosophila
36	3818.5	38.1	1548	2	Q7KTD0_DROME	Q7KTD0 drosophila
37	3813.5	38.1	1548	2	Q7KTC8_DROME	Q7KTC8 drosophila
38	3812.5	38.1	1548	2	Q7KTC4_DROME	Q7KTC4 drosophila
39	3807.5	38.0	1548	2	Q7KTC9_DROME	Q7KTC9 drosophila
40	3793.5	37.9	1548	2	Q7KTC7_DROME	Q7KTC7 drosophila
41	3790	37.8	1549	2	Q9VK56_DROME	Q9VK56 drosophila
42	3771.5	37.7	1548	2	Q7KTC5_DROME	Q7KTC5 drosophila
43	3769.5	37.6	1548	2	Q7KTC6_DROME	Q7KTC6 drosophila
44	3741	37.4	1564	2	08QC98_RAJER	08QC98 raja erinac
45	3738.5	37.3	1567	2	Q6PH26_BRARE	Q6PH26 brachydanio

ALIGNMENTS

RESULT 1
MRP1_HUMAN STANDARD; PRT; 1531 AA.
ID MRP1_HUMAN
AC P33527; 014819; P78419;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Multidrug resistance-associated protein 1 (ATP-binding cassette sub-
family C member 1).
GN Name=ABCC1; Synonyms=MRP, MRP1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93088080; PubMed=1360704;
RA Cole S.P.C., Bhargava G., Gerlach J.H., Mackie J.E., Grant C.E.,
Almquist K.C., Stewart A.J., Kurz E.U., Duncan A.M.V., Deeley R.G.;
RT "Overexpression of a transporter gene in a multidrug-resistant human
lung cancer cell line.";
RL Science 258:1650-1654 (1992).
RN [2]
RP SEQUENCE REVISION.
RX MEDLINE=93262415; PubMed=8098549;
RA Cole S.P.C., Deeley R.G.;
RT "Multidrug resistance-associated protein: sequence correction.";
RL Science 260:879-879 (1993).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96251691; PubMed=8649356;
RA Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P.C.,
Deeley R.G.;
RT "Structure and expression of the messenger RNA encoding the murine
multidrug resistance protein, an ATP-binding cassette transporter.";
RL Mol. Pharmacol. 49:962-971 (1996).
RN [4]
RP NUCLEOTIDE SEQUENCE, AND ALTERNATIVE SPLICING.
RX MEDLINE=8808927; PubMed=9344662; DOI=10.1006/geno.1999.5927;
RA Grant C.E., Kurz E.U., Cole S.P.C., Deeley R.G.;
RT "Analysis of the intron-exon organization of the human multidrug-
resistance protein gene (MRP) and alternative splicing of its mRNA.";
RL Genomics 45:368-378 (1997).
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA] OF 1131-1531.
RX MEDLINE=99425270; PubMed=10493829; DOI=10.1006/geno.1999.5927;
RA Lottus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,
Fuhrmann U., Mason T., Crosby M.L., Barnstead M., Cronin L.,
Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
RT "Genome duplications and other features in 12 Mb of DNA sequence from
human chromosome 16p and 16q.";
RL Genomics 60:295-308 (1999).

RN [6]
 RP TOPOLOGY, AND CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=9744425; PubMed=9295302; DOI=10.1074/jbc.272.38.23623;
 RA Hifner D.R., Almqvist K.C., Leslie E.M., Gerlach J.H., Grant C.E.,
 RA Deley R.G., Cole S.P.C.;
 RT "Membrane topology of the multidrug resistance protein (MRP). A study
 of glycosylation-site mutants reveals an extracytosolic NH2
 terminus."
 RL J. Biol. Chem. 272:23623-23630(1997).
 RN [7]
 RP TOPOLOGY.
 RX MEDLINE=97476249; PubMed=9334225; DOI=10.1074/jbc.272.42.26479;
 RA Kaet C., Gros P.;
 RT "Topology mapping of the amino-terminal half of multidrug resistance-
 associated protein by epitope insertion and immunofluorescence."
 RL J. Biol. Chem. 272:26479-26487(1997).
 RN [8]
 RP TOPOLOGY.
 RX MEDLINE=96153110; PubMed=9485377; DOI=10.1021/bi972332v;
 RA Kaet C., Gros P.;
 RT "Epitope insertion favors a six transmembrane domain model for the
 carboxy-terminal portion of the multidrug resistance-associated
 protein."
 RL Biochemistry 37:2305-2313(1998).
 RN [9]
 RP MUTAGENESIS OF ASP-792; ASP-793; LYS-1333 AND 1454-ASP-1455.
 RX MEDLINE=21362977; PubMed=11469806; DOI=10.1006/abbi.2001.2441;
 RA Cui L., Hou Y.-X., Riordan J.R., Chang X.-B.;
 RT "Mutations of the Walker B motif in the first nucleotide binding
 domain of multidrug resistance protein MRP1 prevent conformational
 maturation."
 RL Arch. Biochem. Biophys. 392:153-161(2001).
 RN [10]
 RP MUTAGENESIS OF TRP-1246.
 RX MEDLINE=21238219; PubMed=11278867; DOI=10.1074/jbc.M011246200;
 RA Ito K., Olsen S.L., Qiu W., Deeley R.G., Cole S.P.C.;
 RT "Mutation of a single conserved tryptophan in multidrug resistance
 protein 1 (MRP1/ABCC1) results in loss of drug resistance and
 selective loss of organic anion transport."
 RL J. Biol. Chem. 276:15616-15624(2001).
 RN [11]
 RP VARIANTS GLN-633 AND VAL-671.
 RX MEDLINE=20283940; PubMed=10835642; DOI=10.1038/76102;
 RA Le Saux O., Urban Z., Tschuch C., Csizsar K., Bacchelli B.,
 RA Quaglini D., Pasquali-Ronchetti I., Pope F.M., Richards A., Terry S.,
 RA Becroft L., de Paep A., Boyd C.D.;
 RT "Mutations in a gene encoding an ABC transporter cause pseudoxanthoma
 elasticum."
 RL Nat. Genet. 25:223-227(2000).
 RN [12]
 RP VARIANT VAL-671.
 RX MEDLINE=20283940; PubMed=10811882; DOI=10.1073/pnas.100041297;
 RA Ringpfeil F., Lehwahl M.G., Cristiano A.M., Uitto J.;
 RT "Pseudoxanthoma elasticum: mutations in the MRP6 gene encoding a
 transmembrane ATP-binding cassette (ABC) transporter."
 RL Proc. Natl. Acad. Sci. U.S.A. 97:6001-6006(2000).
 RN [13]
 RP VARIANTS SER-433, AND CHARACTERIZATION OF VARIANT VAL-671.
 RX MEDLINE=21576494; PubMed=11721885; DOI=10.1007/s100380170017;
 RA Conrad S., Kaufmann H.-M., Ito K., Deeley R.G., Cole S.P.C.,
 RA Schrenk D.;
 RT "Identification of human multidrug resistance protein 1 (MRP1)
 mutations and characterization of a G671V substitution."
 RL J. Hum. Genet. 46:656-663(2001).
 RN [14]
 RP VARIANTS MET-117 AND LEU-1512.
 RX MEDLINE=20579883; PubMed=11139250;
 DOI=10.1002/1098-1004(2001)17:1<74::AID-HUMU14>3.0.CO;2-F;
 RA Perdu J., Germain D.P.;
 RT "Identification of novel polymorphisms in the pms and MRP1 (ABCC1)
 genes at locus 16p13.1 and exclusion of both genes as responsible for
 pseudoxanthoma elasticum."
 RL Hum. Mutat. 17:74-75(2001).

RN [15]
 RP VARIANTS SER-43; ILE-73; GLN-723 AND GLN-1058.
 RX MEDLINE=21163848; PubMed=11266082;
 DOI=10.1097/00008571-200103000-00008;
 RA Ito S., Ito I., Tanabe M., Suzuki A., Higuchi S., Otsubo K.;
 RT "Polymorphism of the ABC transporter genes, MRP1, MRP2 and MRP3/ABCC1,
 in healthy Japanese subjects."
 RL Pharmacogenetics 11:175-184(2001).
 CC -1- FUNCTION: May participate directly in the active transport of
 drugs into subcellular organelles or influence drug distribution
 indirectly.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 Event-Alternative splicing; Named isoforms=8;
 Comment-Additional isoforms seem to exist. Experimental
 confirmation may be lacking for some isoforms;
 Name=Alllexons;
 CC IsoId=P33527-1; Sequence=Displayed;
 CC Name=Delexon-17;
 CC IsoId=P33527-2; Sequence=VSP_000037;
 CC Name=Delexon-18;
 CC IsoId=P33527-3; Sequence=VSP_000038;
 CC Name=Delexon-30;
 CC IsoId=P33527-4; Sequence=VSP_000039;
 CC Name=Delexon-17-18;
 CC IsoId=P33527-5; Sequence=VSP_000037, VSP_000038;
 CC Name=Delexon-17-30;
 CC IsoId=P33527-6; Sequence=VSP_000037, VSP_000039;
 CC Name=Delexon-18-30;
 CC IsoId=P33527-7; Sequence=VSP_000038, VSP_000039;
 CC Name=Delexon-17-18-30;
 CC IsoId=P33527-8; Sequence=VSP_000037, VSP_000038, VSP_000039;
 CC -1- TISSUE SPECIFICITY: Lung, testis and peripheral blood mononuclear
 cells.
 CC -1- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.
 CC -1- SIMILARITY: Contains 2 ABC transmembrane type-1 domains.
 CC -1- DATABASE: NAME=Atlas Genet. Cytoenet. Oncol. Haematol.;
 WWW="http://www.infobio.gen.fr/services/chromosome/gene/MRP106.html".
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.
 CC -----
 CC EMBL, L05628; AAB4616.1; -, mRNA.
 DR EMBL, AF022853; AAB83983.1; -, Genomic DNA.
 DR EMBL, AF022824; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022825; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022826; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022827; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022828; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022829; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022830; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022831; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022832; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022833; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022834; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022835; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022836; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022837; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022838; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022839; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022840; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022841; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022842; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022843; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022844; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022845; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022846; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022847; AAB83983.1; JOINED; Genomic DNA.
 DR EMBL, AF022848; AAB83983.1; JOINED; Genomic DNA.

DR EMBL; AF022849; AAB83983.1; JOINED; Genomic DNA.
DR EMBL; AF022850; AAB83983.1; JOINED; Genomic DNA.

Query Match 78.5%; Score 7860; DB 1; Length 1531;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALRGFCADGSDPELMDNMTNTSNDFKCFONTLVWPCFYLMACPEFYLYSRH 60
DB 1 MALRGFCADGSDPELMDNMTNTSNDFKCFONTLVWPCFYLMACPEFYLYSRH 60
QY 61 DRGYIQTMTKTKTALGFLMIYCMADLFYSFWEBSRGIFLAEVLSPTLGIITLLA 120
DB 61 DRGYIQTMTKTKTALGFLMIYCMADLFYSFWEBSRGIFLAEVLSPTLGIITLLA 120
QY 121 TPLIOLERRKGVSSGIMLTFMLVYCALAIIASKTMTALKEDAOVDLFRDITFYFYS 180
DB 121 TPLIOLERRKGVSSGIMLTFMLVYCALAIIASKTMTALKEDAOVDLFRDITFYFYS 180
QY 181 LLILOLVLSGFSDBSPLESETIHDNPNCPSSASGFLSRITFWITGLIVRGYRQPLSGSD 240
DB 181 LLILOLVLSGFSDBSPLESETIHDNPNCPSSASGFLSRITFWITGLIVRGYRQPLSGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKWKKCAKTRKQPVKVYSSKDPAPQKSSKVDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKWKKCAKTRKQPVKVYSSKDPAPQKSSKVDANEVEAL 300
QY 301 IVSPKQKPNPDLFKVLTKTPGYFLMSFFPKAIIHLMFSGPOILKLIKFNVDIYAPD 360
DB 301 IVSPKQKPNPDLFKVLTKTPGYFLMSFFPKAIIHLMFSGPOILKLIKFNVDIYAPD 360
QY 361 MOGYFYVLLFVPAACIOTLVLAHOYFHTCFVSGMRKTAIVGAYVRKALVITNSARKSSTV 420
DB 361 MOGYFYVLLFVPAACIOTLVLAHOYFHTCFVSGMRKTAIVGAYVRKALVITNSARKSSTV 420
QY 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOVLIILYLLMLTGPSVLAGVAVMLVMPVN 480
DB 421 GEIVNLMSVDAQRFMDLATYINMIWSAPLOVLIILYLLMLTGPSVLAGVAVMLVMPVN 480
QY 481 AVNANKTKTYQVAHMSKDNRIKLMNEILINGIKVLYAMELAFKDKVLAIRBELKVLK 540
DB 481 AVNANKTKTYQVAHMSKDNRIKLMNEILINGIKVLYAMELAFKDKVLAIRBELKVLK 540
QY 541 KSAVLSAVGFTWCPPELVALCTPAAVYVITDENNILLADCTAFATLPMILSPPTNLP 600
DB 541 KSAVLSAVGFTWCPPELVALCTPAAVYVITDENNILLADCTAFATLPMILSPPTNLP 600
QY 601 MVISIVQASVSLKRLRIFLSHELEBDSIERRPVKDGGTNSITVYNAFTWARSDDPT 660
DB 601 MVISIVQASVSLKRLRIFLSHELEBDSIERRPVKDGGTNSITVYNAFTWARSDDPT 660
QY 661 LMGITTSIPREGALVAVVGQVCGCKSSLSALLAEMDKVEGHVAIKGSVAVVPOQAMTQND 720
DB 661 LMGITTSIPREGALVAVVGQVCGCKSSLSALLAEMDKVEGHVAIKGSVAVVPOQAMTQND 720
QY 721 SLRENLIFGQULEEPPYRSYIOACALLPDLIELIPSGRTEITGKGVNLSSGQQRVSILAR 780
DB 721 SLRENLIFGQULEEPPYRSYIOACALLPDLIELIPSGRTEITGKGVNLSSGQQRVSILAR 780
QY 781 AAVSNADIVLPDPLSAVDHVGKILPENYIGPKMKNTKRIILVTHSMGYLPOVDYIIV 840
DB 781 AAVSNADIVLPDPLSAVDHVGKILPENYIGPKMKNTKRIILVTHSMGYLPOVDYIIV 840
QY 841 MSGGKISEMGSYOBELARDAFABFLTYASTROBDAENGVTVSGPCKEAKQKEMNGM 900
DB 841 MSGGKISEMGSYOBELARDAFABFLTYASTROBDAENGVTVSGPCKEAKQKEMNGM 900
QY 901 LVYDSAGKOLQROULSSSSSGDISRHNSITAEIQKAEKKEETWKLMEADKQGTGVKL 960
DB 901 LVYDSAGKOLQROULSSSSSGDISRHNSITAEIQKAEKKEETWKLMEADKQGTGVKL 960
QY 961 SVYWDYKAIKGLFISFLIFLFCMNHVSALASNYWLSLMTDDPIVNGTOEHTKVLRSVYG 1020
DB 961 SVYWDYKAIKGLFISFLIFLFCMNHVSALASNYWLSLMTDDPIVNGTOEHTKVLRSVYG 1020

DB 961 SVYWDYKAIKGLFISFLIFLFCMNHVSALASNYWLSLMTDDPIVNGTOEHTKVLRSVYG 1020
QY 1021 ALGISQIAIFGYSMAVSTIGIILASRCLHVDLHSLIRSMSPFEERPSGNLVNREKEL 1080
DB 1021 ALGISQIAIFGYSMAVSTIGIILASRCLHVDLHSLIRSMSPFEERPSGNLVNREKEL 1080
QY 1081 DTVDSMIPBVIKMFMSLFNVIGACIYIILATPILAIIPPLGIYFYVORFYVASSROL 1140
DB 1081 DTVDSMIPBVIKMFMSLFNVIGACIYIILATPILAIIPPLGIYFYVORFYVASSROL 1140
QY 1141 KRLSVSRSPVSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENQAXYPSIVANWLA 1200
DB 1141 KRLSVSRSPVSHFNETLLGVSVIRAFEBQERFIHQSDLKVDENQAXYPSIVANWLA 1200
QY 1201 VRLCVCNCIVLPALPAVISRHSLSAGLVGSVSLQYTTIYMLNLRVRSSEMETIYA 1260
DB 1201 VRLCVCNCIVLPALPAVISRHSLSAGLVGSVSLQYTTIYMLNLRVRSSEMETIYA 1260
QY 1261 VERLKEYSETEKAPMOIQETAPSSWPQVGRVFPNRYCLRYREDLDFVLRHINVTING 1320
DB 1261 VERLKEYSETEKAPMOIQETAPSSWPQVGRVFPNRYCLRYREDLDFVLRHINVTING 1320
QY 1321 EKVGIYGRTGAGKSSLTGLFRINESAGEI11IDGINIAKIGLHDLRFK1T11PODPVL 1380
DB 1321 EKVGIYGRTGAGKSSLTGLFRINESAGEI11IDGINIAKIGLHDLRFK1T11PODPVL 1380
QY 1381 SGLSRNMLDPFGQYSBEWTSLELAHLKDFVALPDKLDHCEAGGEMISVQORQVLCL 1440
DB 1381 SGLSRNMLDPFGQYSBEWTSLELAHLKDFVALPDKLDHCEAGGEMISVQORQVLCL 1440
QY 1441 ARALLKTKTLVLDENPAVADLFTDLDLQSTIRTOBEDCTVLTIARLNTIMOTRYIVL 1500
DB 1441 ARALLKTKTLVLDENPAVADLFTDLDLQSTIRTOBEDCTVLTIARLNTIMOTRYIVL 1500
QY 1501 DKGEIQEYGA PSDLLQORGLFYSMAKDAGIV 1531
DB 1501 DKGEIQEYGA PSDLLQORGLFYSMAKDAGIV 1531

RESULT 2
Q9UD99 HUMAN
ID Q9UD99 HUMAN PRELIMINARY; PRT; 1515 AA.
AC Q9UD99;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Multidrug resistance protein (Fragment).
GN Name=MRP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;
RA Grant C.B., Kurz E.U., Cole S.P.C., Deeley R.G.;
RT Analysis of the inticon-exon organization of the human multidrug-
resistance protein gene (MRP) and alternative splicing of its mRNA.";
RL Genomics 45:368-378(1997).
CC - - Similarity: Belongs to the ABC transporter family.
DR EMBL; AF022853; AAB83979.1; -; Genomic DNA.
DR EMBL; AF022827; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022828; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022829; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022831; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022833; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022835; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022837; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022839; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022841; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022850; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022849; AAB83979.1; JOINED; Genomic DNA.

DR EMBL; AF022848; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022847; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022846; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022845; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022844; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022843; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022842; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022841; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022840; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022838; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022836; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022834; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022832; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022826; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022825; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022824; AAB83979.1; JOINED; Genomic DNA.
DR EMBL; AF022830; AAB83979.1; JOINED; Genomic DNA.
DR HSSP; P08716; 1MT0.
DR Ensembl; ENSG00000103222; Homo sapiens.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR011527; ABC membrane 1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA_2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS00929; ABC_TMIF_2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS01071; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR KATP-binding; Nucleotide-binding.
KW NON_TER 1
FT SEQUENCE 1515 AA; 169853 MW; DC8552817C439FE CRC64;
SQ
Query March 77.6%; Score 7769; DB 2; Length 1515;
Best local Similarity 100.0%; Pred. No. 0;
Matches 1515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 377 QTLVHOYFHI CFSVGNRIKTAIVIGAYRKALVITNSARKSVTGEIVNLSMDAQRPM 436
DB 361 QTLVHOYFHI CFSVGNRIKTAIVIGAYRKALVITNSARKSVTGEIVNLSMDAQRPM 420
QY 437 LATYINNTNSAPLOVILATLVLTWNLGPSYLAGAVNVLVPPVNAVNAAMTKTYOVAMHK 496
DB 421 LATYINNTNSAPLOVILATLVLTWNLGPSYLAGAVNVLVPPVNAVNAAMTKTYOVAMHK 480
QY 497 SKDRIKLMNEILNGIKVLYLKYAMELAFKDKVLAIROBEELVKLKSAYLSAVGTFTWCT 556
DB 481 SKDRIKLMNEILNGIKVLYLKYAMELAFKDKVLAIROBEELVKLKSAYLSAVGTFTWCT 540
QY 557 PFLVALCTFAVYVITDENNNILDAQTAFVSIALFNILFPLNILPMVSISSVQASVSLKRL 616
DB 541 PFLVALCTFAVYVITDENNNILDAQTAFVSIALFNILFPLNILPMVSISSVQASVSLKRL 600
QY 617 RIFLSHELEPDSIEREPVVDGGTNSITYRNAFTMARSPPTLNGITTSIPFGALVAV 676
DB 601 RIFLSHELEPDSIEREPVVDGGTNSITYRNAFTMARSPPTLNGITTSIPFGALVAV 660
QY 677 VGQVCGCKSSILSLALEMDKVEGHVAIKGSVAIVPQAMVIONDSLRENILFGQLEBPY 736
DB 661 VGQVCGCKSSILSLALEMDKVEGHVAIKGSVAIVPQAMVIONDSLRENILFGQLEBPY 720
QY 737 YRSVIOACALLPDLEILPSGDRTEIGEGVNLGGQKQSVSLARAVYSNADIVYFDDPLS 796
DB 721 YRSVIOACALLPDLEILPSGDRTEIGEGVNLGGQKQSVSLARAVYSNADIVYFDDPLS 780
QY 797 AYDAHVGKHI PENYIGRKMLKNTRIILYHSMGYLQOVNVIYMSGGKISEMSYQELL 856
DB 781 AYDAHVGKHI PENYIGRKMLKNTRIILYHSMGYLQOVNVIYMSGGKISEMSYQELL 840
QY 857 ARDGAFAEFLRTYASTOEODAEENGVTGVSQPEKAKOMENGMVYDSDAKQOLRODSS 916
DB 841 ARDGAFAEFLRTYASTOEODAEENGVTGVSQPEKAKOMENGMVYDSDAKQOLRODSS 900
QY 917 SSSYSGDISRRHNSYAEIOKAKEKETWKLMEADKAQTQGVKLSYVYWDYKAIGLFISF 976
DB 901 SSSYSGDISRRHNSYAEIOKAKEKETWKLMEADKAQTQGVKLSYVYWDYKAIGLFISF 960
QY 977 LSTFLMCHNSALASNYWLSLMTDDPIYNGTOHTKTVRLSYGALGISQIAIFGYSMA 1036
DB 961 LSTFLMCHNSALASNYWLSLMTDDPIYNGTOHTKTVRLSYGALGISQIAIFGYSMA 1020
QY 1037 VSIGIILASRCLATVDLHSLIRSPMSFEPTPSGNLVNRFSEKLDYDMSIPEVIKPMFG 1096
DB 1021 VSIGIILASRCLATVDLHSLIRSPMSFEPTPSGNLVNRFSEKLDYDMSIPEVIKPMFG 1080
QY 1097 SLFNVIGACIVILLATPIAIIIPPLGIYFFVQRFVYASSRQKRLSEVSRSVPYSHFN 1156
DB 1081 SLFNVIGACIVILLATPIAIIIPPLGIYFFVQRFVYASSRQKRLSEVSRSVPYSHFN 1140
QY 1157 ETLIGSVYTRAFEBQERFIHOSDLKVDENQAXIYPSIYANRWLAVRLECVCNCLVPAAL 1216
DB 1141 ETLIGSVYTRAFEBQERFIHOSDLKVDENQAXIYPSIYANRWLAVRLECVCNCLVPAAL 1200
QY 1217 FAVISRHSLSAGVIGLSVSYSLQVTTYLNMVVRMSSEMETNIYAVERLKESETKEKAPM 1276
DB 1201 FAVISRHSLSAGVIGLSVSYSLQVTTYLNMVVRMSSEMETNIYAVERLKESETKEKAPM 1260
QY 1277 QIOETAPSSWPQVGRVEFRNYCLRYBEDLDVLRHINVTINGEKVGI VGRTAGKSSL 1336
DB 1261 QIOETAPSSWPQVGRVEFRNYCLRYBEDLDVLRHINVTINGEKVGI VGRTAGKSSL 1320
QY 1337 TLGLFRINESAEBGIIIDGINIAKIGLHDLRFKTIITIPQDVLPSGSIJMMULDPFSQYSD 1396
DB 1321 TLGLFRINESAEBGIIIDGINIAKIGLHDLRFKTIITIPQDVLPSGSIJMMULDPFSQYSD 1380
QY 1397 EEWMTSLELAHKQPFVSLPDKLDHECAGEGENISVVGOROLVCLARALRRTKTLIVDEA 1456
DB 1381 EEWMTSLELAHKQPFVSLPDKLDHECAGEGENISVVGOROLVCLARALRRTKTLIVDEA 1440
QY 1457 TAAVDLEFDDLLIGSTIRTOFEDCTVLITAHRLANTIMDTRVIVLDKGIQIEYGAPSDLLQ 1516

DB 1441 TAAVLEDTDLIGTIRTFQEDCTVLTATHRANTIMOTRYIVLDBKGIQGYGAPSDLLQ 1500
 QY 1517 QRGIFYMAKXAGIV 1531
 DB 1501 QRGIFYMAKXAGIV 1515

RESULT 3

0864R9_MACFA
 ID 0864R9_MACFA PRELIMINARY; PRT; 1531 AA.
 AC 0864R9;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Multidrug resistance protein 1B.
 GN Name=MRP1.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 OC Cercopitheciidae; Cercopitheciinae; Macaca.
 OC NCBI_TaxID=9541;
 RX NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22544876; PubMed=12657726;
 RA Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,
 RA Dantzig A.H., Perry W.L.;
 RT "Cloning and functional characterization of the multidrug resistance-
 RT associated protein (MRP1/ABCC1) from the cynomolgus monkey.";
 RL Mol. Cancer Ther. 2:307-316(2003).
 RN [2]

NUCLEOTIDE SEQUENCE.

RA Perry W.L., Godinot N.;
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the ABC transporter family.
 DR EMBL; AY146673; AAN65349.1; -; mRNA.
 DR HSSP; P08716; IMTO.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005524; P:ATP binding; IEA.
 DR GO; GO:0016887; P:ATPase activity; IEA.
 DR GO; GO:0042666; F:ATPase activity, coupled to transmembrane m. . .; IEA.
 DR GO; GO:0000166; F:nucleotide binding; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR011527; ABC_membrane_1.
 DR InterPro; IPR001140; ABC_TM_transp.
 DR InterPro; IPR003439; ABC_transp_like.
 DR InterPro; IPR005292; NRP_assoc.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00664; ABC_membrane; 2.
 DR Pfam; PF00005; ABC_tran; 2.
 DR Prodom; PD000006; ABC_transporter; 2.
 DR SMART; SM00382; AAA; 2.
 DR TIGRFAMs; TIGR00957; NRP_assoc_pro; 1.
 DR PROSITE; PS50929; ABC_TMIF_2.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
 DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 KM ATP-binding; Nucleotide-binding; Repeat.
 SQ SEQUENCE 1531 AA; 171659 MW; 1AE788EFDF9EF459 CRC64;

Query Match 76.8%; Score 7695; DB 2; Length 1531;
 Best Local Similarity 97.5%; Pred. No. 0;
 Matches 1493; Conservative 20; Mismatches 18; Indels 0; Gaps 0;

QY 1 MALRGFCADGSDPLMDMNTVNTSNPDTKCFONTVLVWVPCFYLMACFPFYFLYLSRH 60
 DB 1 MALRGFCADGSDPLMDMNTVNTSNPDTKCFONTVLVWVPCFYLMACFPFYFLYLSRH 60
 QY 61 DRGTYQWTPLNKTKTALGFLIMIVCMADLFYSFWEBSRGIFLADVPVLSPTLLGITTLLA 120
 DB 61 DRGTYQWTPLNKTKTALGFLIMIVCMADLFYSFWEBSRGIFLADVPVLSPTLLGITTLLA 120

QY 121 TFLIQLERRKGVSSGIMLTFWLVAVCALAIILRSKIMTALKEDAQOVLFRDITFVYYS 180
 DB 121 TFLIQLERRKGVSSGIMLTFWLVAVCALAIILRSKIMTALKEDAQOVLFRDITFVYYS 180
 QY 181 LLLIQLVLSGFSRDRPLFSETIHDNDCPSBSAFLSRIITFWMITGILVRGROPLRSGSD 240
 DB 181 LLLIQLVLSGFSRDRPLFSETIHDNDCPSBSAFLSRIITFWMITGILVRGROPLRSGSD 240
 QY 241 LMSLNKEDTSBQVVPVYKWKWKGCATKROKPVKVVYSSKDPAPKSSSKVDANEVEAL 300
 DB 241 LMSLNKEDTSBQVVPVYKWKWKGCATKROKPVKVVYSSKDPAPKSSSKVDANEVEAL 300
 QY 301 IVKSPQKEMPSLSEKVLKTFGGPYFLMSFFPKAHDLMFSGPOLLKLLKFVNDTAPD 360
 DB 301 IVKSPQKEMPSLSEKVLKTFGGPYFLMSFFPKAHDLMFSGPOLLKLLKFVNDTAPD 360
 QY 361 WQGYFTVLLFVTAQLOTLVHQYFHLCPVSGMRKTAIVIGAVRKALVITNSARKSTV 420
 DB 361 WQGYFTVLLFVTAQLOTLVHQYFHLCPVSGMRKTAIVIGAVRKALVITNSARKSTV 420
 QY 421 GEIVNLMVSVDQRPMDLATTINMTWSPAPLOYITLALYLMNLGPSVLAGVAVMYLMPVN 480
 DB 421 GEIVNLMVSVDQRPMDLATTINMTWSPAPLOYITLALYLMNLGPSVLAGVAVMYLMPVN 480
 QY 481 AVYAMKTKTYQVAMKSKDKNRKILMNLINGIKVLKIYANELAKDKVLAIRBELKVLK 540
 DB 481 AVYAMKTKTYQVAMKSKDKNRKILMNLINGIKVLKIYANELAKDKVLAIRBELKVLK 540
 QY 541 KSAVLISAVGTFTWCTPFLVALCTFAVYVITIDENNILDAOTAFAFSLALFNILRPILNP 600
 DB 541 KSAVLISAVGTFTWCTPFLVALCTFAVYVITIDENNILDAOTAFAFSLALFNILRPILNP 600
 QY 601 MVISSIVQASVSLKRLRIPLSHELBSPDSEIRRPVKDGDTNSTVYANATFTMARSDPT 660
 DB 601 MVISSIVQASVSLKRLRIPLSHELBSPDSEIRRPVKDGDTNSTVYANATFTMARSDPT 660
 QY 661 LMGTFEIPRGALVAVVGVGCCGKSSLSLLALMEMDVBEHVAKGSAVVPQAMQND 720
 DB 661 LMGTFEIPRGALVAVVGVGCCGKSSLSLLALMEMDVBEHVAKGSAVVPQAMQND 720
 QY 721 SLRENILFGQLEBPYRSYIQAALLPDEILPDSGRTEIGEGVMISGGOKORVSLAR 780
 DB 721 SLRENILFGQLEBPYRSYIQAALLPDEILPDSGRTEIGEGVMISGGOKORVSLAR 780
 QY 781 AVYSNADIYFDDPLSAVDHAVGKHIPENVIQPKMKLKNKTRILVTHSMYSLPQVDYIIV 840
 DB 781 AVYSNADIYFDDPLSAVDHAVGKHIPENVIQPKMKLKNKTRILVTHSMYSLPQVDYIIV 840
 QY 841 MSGKTISEMSYQELLARDGAFAPFLRTASTQEOQAEENGVTGVSQPGKEAKQEMNGM 900
 DB 841 MSGKTISEMSYQELLARDGAFAPFLRTASTQEOQAEENGVTGVSQPGKEAKQEMNGM 900
 QY 901 LYTDSAGKOROLRSSSSSYSGDISRRHNSTAELQKBAKKEFTWKMEADKAOTGVOKL 960
 DB 901 LYTDSAGKOROLRSSSSSYSGDISRRHNSTAELQKBAKKEFTWKMEADKAOTGVOKL 960
 QY 961 SYTWYDVKAKGLFISLFLFMCNHYASALASNYWLSLWTDPIVNGTOEHTKVLRSVYG 1020
 DB 961 SYTWYDVKAKGLFISLFLFMCNHYASALASNYWLSLWTDPIVNGTOEHTKVLRSVYG 1020
 QY 1021 ALGISQGIANFGYSMAVSIIGIILASRCLHVDLHSLIRSPMSFPERPISGNLVNRSEKEL 1080
 DB 1021 ALGISQGIANFGYSMAVSIIGIILASRCLHVDLHSLIRSPMSFPERPISGNLVNRSEKEL 1080
 QY 1081 DTYDSMIPVYIKPMFMSLFVNIACIYIILATPIAIIIPPLGIYFFVORFYAASSROL 1140
 DB 1081 DTYDSMIPVYIKPMFMSLFVNIACIYIILATPIAIIIPPLGIYFFVORFYAASSROL 1140
 QY 1141 KRLBSVSRSPVYGHFNELLGVSVITAFEEQERFIHOSDLKVDENOKAAYPSIYANRWLA 1200
 DB 1141 KRLBSVSRSPVYGHFNELLGVSVITAFEEQERFIHOSDLKVDENOKAAYPSIYANRWLA 1200
 QY 1201 VRLBVCNCLVFAALFAVISRSLSGVLGVSYSLOQYTTIANTLVRRSSSEMETNIVA 1260

Db 1201 VRLCVCAGCIVLFAALFAVIRSHSISAGLVGSYSIQVTTYTNMVLVMSSEMETIVA 1260
Qy 1261 VERLKEVSETEKEAPWMOIQETAPPSWPVOVGEVFRNCLARYEDLDVLRHINTINGG 1320
Db 1261 VERLKEVSETEKEAPWMOIQETAPPSWPVOVGEVFRNCLARYEDLDVLRHINTINGG 1320
Qy 1321 EKVGVIGRTGAGKSSSLTGLFRINSAAGEI11DGINIAK1GLHDLRFKRT11IPDPVLF 1380
Db 1321 EKVGVIGRTGAGKSSSLTGLFRINSAAGEI11DGINIAK1GLHDLRFKRT11IPDPVLF 1380
Qy 1381 SGSLRMLNDPFSQYSDEEVTWLSLELAHKDFYSALPKDLHRCAGGENLSVGQOLVCL 1440
Db 1381 SGSLRMLNDPFSQYSDEEVTWLSLELAHKDFYSALPKDLHRCAGGENLSVGQOLVCL 1440
Qy 1441 ARALLRKRTILVLDATAVADLETDDLIOSTIRTOFEDCTVLT1AHLNT1MDYTRV1VL 1500
Db 1441 ARALLRKRTILVLDATAVADLETDDLIOSTIRTOFEDCTVLT1AHLNT1MDYTRV1VL 1500
Qy 1501 DKGEIOEYGAPSDLLQQRGLFYSMAKDAGLV 1531
Db 1501 DKGEIOEYGAPSDLLQQRGLFYSMAKDAGLV 1531

RESULT 4

Q864S0_MACFA PRELIMINARY; PRF; 1531 AA.
ID Q864S0; AC Q864S0;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Multidrug resistance protein 1A.
OS Name=MPI;
GN Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecoidea; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RA Godinot N., Iversen P.W., Tabas L., Xia X., Williams D.C.,
RA Dantzig A.H., Perry W.L.;
RT "Cloning and functional characterization of the multidrug resistance-
RT associated protein (MRP1/ABCC1) from the cynomolgus monkey.",
RL Mol. Cancer Ther. 2:307-316(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;
RA Perry W.L., Godinot N.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY146672; AA05346.1; -, mRNA.
DR HSSP; P08716; IMTO.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR01527; ABC_ATPase.
DR InterPro; IPR01140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane_2.
DR Pfam; PF00005; ABC_tran_2.
DR ProDom; PD000006; ABC_transporter_2.
DR SMART; SM00382; AAA_2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro_1.
DR PROSITE; PS50929; ABC_TMIF_2.

DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS0107; PROTEIN KINASE_ATP; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SEQUENCE 1531 AA; 171701 MW; 1687128CC2DB89 CRC64;

Query Match 76.8%; Score 7689; DB 2; Length 1531;
Best Local Similarity 97.6%; Pred. No. 0;
Matches 1494; Conservative 18; Mismatches 19; Indels 0; Gaps 0;

1 MALRGPCSDGSDMLMDMNTWMTNPNDFPKCPONTLVWPCGYLWACPFYFLYSRH 60
1 MALRGPCSDGSDMLMDMNTWMTNPNDFPKCPONTLVWPCGYLWACPFYFLYSRH 60
Qy 61 DRGYIOMTPUNKRTALGFLIMICWADLFYSFMERSRGIFLAVFLVSPYLLGITTLA 120
Db 61 DRGYIOMTPUNKRTALGFLIMICWADLFYSFMERSRGIFLAVFLVSPYLLGITTLA 120
Qy 121 TFLIQLERRKGVSSGIMLTFWLVALVCAALIRSK1MTALKEDAOVDLFRDITFYVES 180
Db 121 TFLIQLERRKGVSSGIMLTFWLVALVCAALIRSK1MTALKEDAOVDLFRDITFYVES 180
Qy 181 LLLIQVLVSCFSDSPFSETIHDNPNCPSSASFLSRITFWMTTGLIVRGYRQPLEGSD 240
Db 181 LVL1QLVSCFSDSPFSETIHDNPNCPSSASFLSRITFWMTTGLIVRGYRQPLEGSD 240
Qy 241 LMSLINKEDTSBOVVPVLVYKWKKECAKTRKQPVYVYSSKDPAPKSSKYDANEVEAL 300
Db 241 LMSLINKEDTSBOVVPVLVYKWKKECAKTRKQPVYVYSSKDPAPKSSKYDANEVEAL 300
Qy 301 IVKSPQKEMNDLFRVLYKTFGPYFLMSFPFKAIHDLMSGPOILKLLIKFVNDYKAPD 360
Db 301 IVKSPQKEMNDLFRVLYKTFGPYFLMSFPFKAIHDLMSGPOILKLLIKFVNDYKAPD 360
Qy 361 WQGFYTYVLAFTVACLOTIVLHQYFHI CFVSGMGIKTAIVGAVYRKALVITNARKSTV 420
Db 361 WQGFYTYVLAFTVACLOTIVLHQYFHI CFVSGMGIKTAIVGAVYRKALVITNARKSTV 420
Qy 421 GEIVNLSVDAQRFMDLATVYINMISAPLOVITLALYLLMLLGSVLAVGAVMYLMEVFN 480
Db 421 GEIVNLSVDAQRFMDLATVYINMISAPLOVITLALYLLMLLGSVLAVGAVMYLMEVFN 480
Qy 481 AVMAKTKTYOVAMKSKDNRIKIMNELINGIKVLYAMELAFKDYALAIROBELKVLK 540
Db 481 AVMAKTKTYOVAMKSKDNRIKIMNELINGIKVLYAMELAFKDYALAIROBELKVLK 540
Qy 541 KSATLSAVGTFTWCTPFLVALCTFAYVYITDENNIIDAQAFPSLALFNILRFPNTLP 600
Db 541 KSATLSAVGTFTWCTPFLVALCTFAYVYITDENNIIDAQAFPSLALFNILRFPNTLP 600
Qy 601 MVISIYQASVSLRLRIFLSHELEPDSIERRPVKQGGTNSITVRNATFTMARSDPT 660
Db 601 MVISIYQASVSLRLRIFLSHELEPDSIERRPVKQGGTNSITVRNATFTMARSDPT 660
Qy 661 LNGITFSIPREGALVAVVGVQVCGKSSLLSALLAEMDVKEGHVAKGSVAVVYPOQAM1QND 720
Db 661 LNGITFSIPREGALVAVVGVQVCGKSSLLSALLAEMDVKEGHVAKGSVAVVYPOQAM1QND 720
Qy 721 SLRENILPFGCOLPEPYRVSVAQACALLPDLEILPSGRTIEGEGVNLSGQOKRVSLAR 780
Db 721 SLRENILPFGCOLPEPYRVSVAQACALLPDLEILPSGRTIEGEGVNLSGQOKRVSLAR 780
Qy 781 AVYSNADLYLPDDPLASVADAVGHKIFENNVIGPGMLKNKTRILVTHSMSTLPPOVDV1V 840
Db 781 AVYCNADLYLPDDPLASVADAVGHKIFENNVIGPGMLKNKTRILVTHSMSTLPPOVDV1V 840
Qy 841 MSGKISEMSGYOELLARDGAFAEFLLRTYASTEOBDAEENGVTGVSQPGKEAKOMENGM 900
Db 841 MSGKISEMSGYOELLARDGAFAEFLLRTYASTEOBDAEENGVTGVSQPGKEAKOMENGM 900
Qy 901 LVYDSAGKOLQROLSSSSSYSGDISRHNSGTAB1QKAPAKKEETWKLMEADKATQGVKL 960
Db 901 LVYDSAGKOLQROLSSSSSYSGDVSRHNSGTAB1QKAPAKKEETWKLMEADKATQGVKL 960

[illegible]

Dd	301	LYKTFGEFFLMSFFFKALHDMMBGGPOLKLLIFVNDYKAPMDQGFYTYLLEFVACL	360
Qy	377	QTLVLHOFYFHICFVSQMRITKTAIVAGYRKALVITNSARKSSTVGEIVNLMSVDQRFMD	436
Dd	361	QTLVHOYFHCIFCVSGMRITKTAIVAGYRKALVITNSARKSSTVGEIVNLMSVDQRFMD	420
Qy	437	LATYNNMMSAPLOYILALYLLMLNLSBVLAVGVAWMLMBPVNAVMMKTKTYOVAMK	496
Dd	421	LATYNNMMSAPLOYILALYLLMLNLSBVLAVGVAWMLMBPVNAVMMKTKTYOVAMK	480
Qy	497	SKDNRIKIMNELINGIKYUKLYAMELAFQDKYLAIROBELKYUKKSAYLSAVGETYWC	556
Dd	461	SKDNRIKIMNELINGIKYUKLYAMELAFQDKYLAIROBELKYUKKSAYLSAVGETYWC	540
Qy	557	PELVALCFPAVYVYTDENNIIIDAQTAFPVSLAFNLLPREPNTLPMVISIYQASVSLRL	616
Dd	541	PELVALCFPAVYVYTDENNIIIDAQTAFPVSLAFNLLPREPNTLPMVISIYQASVSLRL	600
Qy	617	RIFLSHEELPEPSITERRPYKQGGGINSITVRATATWARSDDPTLANGITFSIPEGALYAV	676
Dd	601	RIFLSHEELPEPSITERRPYKQGGGINSITVRATATWARSDDPTLANGITFSIPEGALYAV	660
Qy	677	VQVQCGKSSLLSALLAEMDKVEGHVAIKGSAAVYPOQAMIONDSIRENIIIFGCLEBPY	736
Dd	661	VQVQCGKSSLLSALLAEMDKVEGHVAIKGSAAVYPOQAMIONDSIRENIIIFGCLEBPY	720
Qy	737	YRSVIOACALLPDLEILPBGDRTEIIGEKGVNLSSGQKORVSLARAVYSNADITYLFDPLPS	796
Dd	721	YRSVIOACALLPDLEILPBGDRTEIIGEKGVNLSSGQKORVSLARAVYSNADITYLFDPLPS	747
Qy	797	AVDAVQKHIIFENVYGPCKMLKNKTRILVYTHMSYLPOVDVITYNSGGKISMSGYOBLL	856
Dd	748	-----KTRILVTHMSYLPVDVITYNSGGKISMSGYOBLL	784
Qy	857	ARDGAFAEFLRTYASTBOEODAEKNGYGVSGSPGKAKENMGMLVYTSAGQOLROJSS	916
Dd	785	ARDGAFAEFLRTYASTBOEODAEKNGYGVSGSPGKAKENMGMLVYTSAGQOLROJSS	844
Qy	917	SSSYSGDISRRHNSYTAELQKAEKKEBTWKMEAPKAQOTGOYKLSYVMDYMAKIGLFIISF	976
Dd	845	SSSYSGDISRRHNSYTAELQKAEKKEBTWKMEAPKAQOTGOYKLSYVMDYMAKIGLFIISF	904
Qy	977	LSIF.PMCMHVSALASNTWLSLMTDDPIYNGIOEHTKYRLSYGALGISOGIAVFGYSMA	1036
Dd	905	LSIF.PMCMHVSALASNTWLSLMTDDPIYNGIOEHTKYRLSYGALGISOGIAVFGYSMA	964
Qy	1037	VSIGGILASRCHVLDLHSILNSPMSFEERTSPSGMLVNRFSKELDTYVSMIPEVIMKFMFG	1096
Dd	965	VSIGGILASRCHVLDLHSILNSPMSFEERTSPSGMLVNRFSKELDTYVSMIPEVIMKFMFG	1022
Qy	1097	SLFNVIIGACTIVILLATPILAAIIPPLGILYFPVQOFYVYASRQOLRLSESVRSPIVSHFN	1156
Dd	1025	SLFNVIIGACTIVILLATPILAAIIPPLGILYFPVQOFYVYASRQOLRLSESVRSPIVSHFN	1088
Qy	1157	ETLLGVSVIRAFEBQERFIHQSDLKVDENOKAYYBSIVANRWLAVRLCEVGNCIYLPAL	1216
Dd	1085	ETLLGVSVIRAFEBQERFIHQSDLKVDENOKAYYBSIVANRWLAVRLCEVGNCIYLPAL	1144
Qy	1217	PAVISRHSISAGLVGSVYSIQYTTYLNMVLRMSSEMETNIVANBRLEKSEYETKEAPW	1276
Dd	1145	PAVISRHSISAGLVGSVYSIQYTTYLNMVLRMSSEMETNIVANBRLEKSEYETKEAPW	1204
Qy	1277	QIOETAPPSMPOVGSVEPRNYCLARYREDLPVLRHINVTINGKEVGIIVGRTGAKGSSL	1336
Dd	1205	QIOETAPPSMPOVGSVEPRNYCLARYREDLPVLRHINVTINGKEVGIIVGRTGAKGSSL	1266
Qy	1337	TLGLFRINMSAGEIITIDGINIAKIGLHDLRFKTIITIPQDPVLFGSGLRMNLDPFQOYSD	1396
Dd	1265	TLGLFRINMSAGEIITIDGINIAKIGLHDLRFKTIITIPQDPVLFGSGLRMNLDPFQOYSD	1332
Qy	1397	EEVWTSLELAHLKDPFVSALPDYKLDHECAGGENLSYGOQVYCLARALLRKTKIIVLDEA	1456
Dd	1335	EEVWTSLELAHLKDPFVSALPDYKLDHECAGGENLSYGOQVYCLARALLRKTKIIVLDEA	1384

QY	1457	1AAVDLEFDLIDICTITRQPEDCVLTITARIKNTIMYTRIVYVDKKEIQEYKAPSDLLQ	1516
Db	1365	1AAVDLEFDLIDICTITRQPEDCVLTITARIKNTIMYTRIVYVDKKEIQEYKAPSDLLQ	1444
QY	1517	ORGLFYSMAKDAGLV	1531
Db	1445	ORGLFYSMAKDAGLV	1459
RESULT 6			
OSUOAO_HUMAN			
ID	OSUOAO_HUMAN	PRELIMINARY;	PRT; 1456 AA.
AC	OSUOAO_		
DT	01-MAY-2000	(TREMBLrel. 13, Created)	
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)	
DT	01-MAR-2004	(TREMBLrel. 26, Last annotation update)	
DE	Multidrug resistance protein (Fragment).		
GN	Name=MRP;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;		
CC	Hom.		
OX	NCBI	TaxID=9606;	
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RX	MEDLINE=98008927; PubMed=9344662; DOI=10.1006/geno.1997.4950;		
RA	Grat C.E., Kurz E.O., Cole S.P.C., Deeley R.G.;		
RT	"Analysis of the intron-exon organisation of the human multidrug		
RT	resistance protein gene (MRP) and alternative splicing of its mRNA."		
CC	Genomics 45:368-378(1997).		
RL	- i - SIMILARITY: Belongs to the ABC transporter family.		
DR	EMBL	AF022853; AAB83981.1;	-7 Genomic DNA.
DR	EMBL	AF022824; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022825; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022826; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022827; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022828; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022829; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022831; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022833; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022835; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022837; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022840; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022841; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022842; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022843; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022844; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022845; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022846; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022847; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022848; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022849; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022850; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022851; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022852; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022853; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022854; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022855; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022856; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022857; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022858; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022859; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022860; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022861; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022862; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022863; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022864; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022865; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022866; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022867; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022868; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022869; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022870; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022871; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022872; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022873; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022874; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022875; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022876; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022877; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022878; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022879; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022880; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022881; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022882; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022883; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022884; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022885; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022886; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022887; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022888; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022889; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022890; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022891; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022892; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022893; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022894; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022895; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022896; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022897; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022898; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022899; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022900; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022901; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022902; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022903; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022904; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022905; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022906; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022907; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022908; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022909; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022910; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022911; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022912; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022913; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022914; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022915; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022916; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022917; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022918; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022919; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022920; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022921; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022922; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022923; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022924; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022925; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022926; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022927; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022928; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022929; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022930; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022931; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022932; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022933; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022934; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022935; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022936; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022937; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022938; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022939; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022940; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022941; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022942; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022943; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022944; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022945; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022946; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022947; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022948; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022949; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022950; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022951; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022952; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022953; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022954; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022955; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022956; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022957; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022958; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022959; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022960; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022961; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022962; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022963; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022964; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022965; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022966; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022967; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022968; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022969; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022970; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022971; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022972; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022973; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022974; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022975; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022976; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022977; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022978; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022979; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022980; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022981; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022982; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022983; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022984; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022985; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022986; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022987; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022988; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022989; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022990; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022991; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022992; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022993; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022994; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022995; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022996; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022997; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022998; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF022999; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023000; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023001; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023002; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023003; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023004; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023005; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023006; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023007; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023008; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023009; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023010; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023011; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023012; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023013; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023014; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023015; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023016; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023017; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023018; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023019; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023020; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023021; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023022; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023023; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023024; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023025; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023026; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023027; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023028; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023029; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023030; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023031; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023032; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023033; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023034; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023035; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023036; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023037; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023038; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023039; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023040; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023041; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023042; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023043; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023044; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023045; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023046; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023047; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023048; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023049; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023050; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023051; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023052; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023053; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023054; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023055; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023056; AAB83981.1;	JOINED; Genomic DNA.
DR	EMBL	AF023057; AAB83981.1;	JOINED; Genomic DNA

DR InterPro: IPR005292; MRP_assoc.
 DR InterPro: IPR000719; Prot_kinase.
 DR Pfam: PF00664; ABC_membrane; 2.
 DR Pfam: PF00005; ABC_tran; 2.
 DR ProDom: PD000006; ABC_transporter; 2.
 DR SMART: SM00382; AAA; 2.
 DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
 DR PROSITE: PS00929; ABC_TMif; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
 DR PROSITE: PS00893; ABC_TRANSPORTER_2; 2.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR ATP-binding; Nucleotide-binding.
 FT NON TER 1
 SQ SEQUENCE 1456 AA; 163233 MW; 8DE8AAB2BC481F2 CRC64;

Query Match 74.1%; Score 7419.5; DB 2; Length 1456;
 Best Local Similarity 96.1%; Pred. No. 0;
 Matches 1456; Conservative 0; Mismatches 0; Indels 59; Gaps 1;

QY 17 DMNVTNTPNDPFTKCFQNTVTVVPCFYLMACFEYFLYLSRHRDGYIQMTPIKTKTA 76
 DB 1 DMNVTNTPNDPFTKCFQNTVTVVPCFYLMACFEYFLYLSRHRDGYIQMTPIKTKTA 60

QY 77 LGFLMTIVCMADLFYSFMERSRGIFLAPVFLVSPFLIGITTLATFLIOLERRKGVSSG 136
 DB 61 LGFLMTIVCMADLFYSFMERSRGIFLAPVFLVSPFLIGITTLATFLIOLERRKGVSSG 120

QY 137 IMTLFWLVVCAALAIIRSKIMTALKEDAQVDLFRDITFYVFSLLIIQLVLCFSRSP 196
 DB 121 IMTLFWLVVCAALAIIRSKIMTALKEDAQVDLFRDITFYVFSLLIIQLVLCFSRSP 180

QY 197 LFSETHDNPCEPSSASFLSRITFMWITGLIVGQVQPLEGSLMNLNEDTSEQVVPV 256
 DB 181 LFSETHDNPCEPSSASFLSRITFMWITGLIVGQVQPLEGSLMNLNEDTSEQVVPV 240

QY 257 LVKNWKECAKTRKQPKVYVSSKDPAPQKSSKYDNEEVEALIVSPKEMNPFLFKV 316
 DB 241 LVKNWKECAKTRKQPKVYVSSKDPAPQKSSKYDNEEVEALIVSPKEMNPFLFKV 300

QY 317 LYTFPGYFLMSFFPKAIDHLMFSGPOLIKLIKFNVDTKAPDMQGYFTVLLFTVACL 376
 DB 301 LYTFPGYFLMSFFPKAIDHLMFSGPOLIKLIKFNVDTKAPDMQGYFTVLLFTVACL 360

QY 377 QTVLHYEHCIFSGGRITAVIGAVYRKALVITNSARKSSTYGEIVNLMSVDAQFMD 436
 DB 361 QTVLHYEHCIFSGGRITAVIGAVYRKALVITNSARKSSTYGEIVNLMSVDAQFMD 420

QY 437 LATYINMWAPLOVIALYILMLNLGSPVLAVAVVWLVAVNAVMAKTKTYQVAHMK 496
 DB 421 LATYINMWAPLOVIALYILMLNLGSPVLAVAVVWLVAVNAVMAKTKTYQVAHMK 480

QY 497 SKONRIKLMEIINGIKVLKLYAMELAFKDKVLAIROBELKVLKKSAYLSAVGTFTVCT 556
 DB 481 SKONRIKLMEIINGIKVLKLYAMELAFKDKVLAIROBELKVLKKSAYLSAVGTFTVCT 540

QY 557 PELVALCTFAVYVITDENNILDQAOTAVSLAFNIIILFPIIIPMTVSSIVQASVSLKL 616
 DB 541 PELVALCTFAVYVITDENNILDQAOTAVSLAFNIIILFPIIIPMTVSSIVQASVSLKL 600

QY 617 RIFLSHEELPDSIERPVDGGGTSITVRNATFTWARSDDPLNLNGITSIEPGALVAV 676
 DB 601 RIFLSHEELPDSIERPVDGGGTSITVRNATFTWARSDDPLNLNGITSIEPGALVAV 660

QY 677 VGVGCGKSSLLSALLAEMDKVEGHVAKSVAVVPOAWIIONDSLRENTILFGQOLEEPY 736
 DB 661 VGVGCGKSSLLSALLAEMDKVEGHVAI----- 688

QY 737 YREVIQACALLPDLLEILPSGDRTEIGRKYNLSSGCKQKORSLARAVYSNDIYLFDPPLS 796
 DB 689 -----KGNLSSGCKQKORSLARAVYSNDIYLFDPPLS 721

QY 797 AYDAHVAKHIFENVIGPKGMLKNKTRILVTHSMSYLPQVAVIIVMSGKISSEMGSYOEL 856
 DB 797 AYDAHVAKHIFENVIGPKGMLKNKTRILVTHSMSYLPQVAVIIVMSGKISSEMGSYOEL 856

DB 722 AYDAHVAKHIFENVIGPKGMLKNKTRILVTHSMSYLPQVAVIIVMSGKISSEMGSYOEL 781
 QY 857 ARQAFAPFLRTYASTBOBDAEENGVTGVSGPCKEKKOMENGLVYDSAGKOLQROLS 916
 DB 782 ARQAFAPFLRTYASTBOBDAEENGVTGVSGPCKEKKOMENGLVYDSAGKOLQROLS 841

QY 917 SSSYSGDISHHNSTAELOKAEAKKEBTWKLMEADKQOTQVXLVYVMDYKAIGLFISF 976
 DB 842 SSSYSGDISHHNSTAELOKAEAKKEBTWKLMEADKQOTQVXLVYVMDYKAIGLFISF 901

QY 977 LSIFLEMCNHSVALASNYMLSLWTDPIVNGTOEHTKVLRSYVGAIGISGIAVFGYMA 1036
 DB 902 LSIFLEMCNHSVALASNYMLSLWTDPIVNGTOEHTKVLRSYVGAIGISGIAVFGYMA 961

QY 1037 VSIIGILASCLHYDLHSLIRSPMSFEPTPSGNLVNRSKELDTYDSMIPEYIKMPMG 1096
 DB 962 VSIIGILASCLHYDLHSLIRSPMSFEPTPSGNLVNRSKELDTYDSMIPEYIKMPMG 1021

QY 1097 SLFNVIGACIVILLATPIAIIIPPLGLIYFPQRFVAVASSROLKLESVSRSPVSHFN 1156
 DB 1022 SLFNVIGACIVILLATPIAIIIPPLGLIYFPQRFVAVASSROLKLESVSRSPVSHFN 1081

QY 1157 ETLGVSVIRAFBOERFIHQSDLKVDENQKAYYPSIVANRWLAVRLECVGNCTVLPAL 1216
 DB 1082 ETLGVSVIRAFBOERFIHQSDLKVDENQKAYYPSIVANRWLAVRLECVGNCTVLPAL 1141

QY 1217 FAVISRHSLSAGLVGSVSLQVTTYLNNLVNMSSEMETNIVAVERLKEYSFEKEAPW 1276
 DB 1142 FAVISRHSLSAGLVGSVSLQVTTYLNNLVNMSSEMETNIVAVERLKEYSFEKEAPW 1201

QY 1277 QIOTAPSSWPQVGRFEFPNVCYLRREDLPVLRIHNTVINGEKVGIYGRGAGSSL 1336
 DB 1202 QIOTAPSSWPQVGRFEFPNVCYLRREDLPVLRIHNTVINGEKVGIYGRGAGSSL 1261

QY 1337 TLGLFRINSEAGEIIIDGINIAKIGHDLRFITIIIPQPVLFSGSLRMLDPFSGYSD 1396
 DB 1262 TLGLFRINSEAGEIIIDGINIAKIGHDLRFITIIIPQPVLFSGSLRMLDPFSGYSD 1321

QY 1397 EEWYTSLELAHLDPVSALPDKLDHCEAGGENLSVGORQLVCLARALLRKTKILVDEA 1456
 DB 1322 EEWYTSLELAHLDPVSALPDKLDHCEAGGENLSVGORQLVCLARALLRKTKILVDEA 1381

QY 1457 TAAVDETDLIOSTRTQREDCTVLTARLNTIDYFVILYDGEIOEYCAPSLIQ 1516
 DB 1382 TAAVDETDLIOSTRTQREDCTVLTARLNTIDYFVILYDGEIOEYCAPSLIQ 1441

QY 1517 QRGIFYSMADAGLV 1531
 DB 1442 QRGIFYSMADAGLV 1456

RESULT 7
 Q6UR05 CANFA
 ID Q6UR05 CANFA PRELIMINARY; PRT; 1531 AA.
 AC Q6UR05;
 DT 05-JUL-2004 (TReMBLrel. 27, Created)
 DT 05-JUL-2004 (TReMBLrel. 27, last sequence update)
 DT 05-JUL-2004 (TReMBLrel. 27, last annotation update)
 DE Multidrug resistance-associated protein 1.
 GN Name=MRP1;
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
 OC Canis.
 NC NCBI_TaxID=9615;
 RN [1]
 RP NCBIOTIDE SEQUENCE.
 RX MEDLINE=22404678; PubMed=12516967;
 RA Ma L., Pratt S.B., Cao J., Dantzig A.H., Moore R.E., Slapak C.A.;
 RT "Identification and characterization of the canine multidrug
 resistance-associated protein.";
 RL Mol. Cancer Ther. 1:1335-1342(2002).
 RN [2]

RP NUCLEOTIDE SEQUENCE.
RA Ma L., Prate S.E., Cao J., Danczig A.H., Moore R.E., Slapac C.A.;
RL Submitted (Aug-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SMILARITY: Belongs to the ABC transporter family.
DR EMBL: AY631728; AAC03148.1; -; mRNA.
DR EMBL: ENSCAFG00000018208; Cantis familiaris.
DR GO: GO:0016021; C:Integral to membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0016887; F:ATPase activity; IEA.
DR GO: GO:0042656; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:0000166; F:nucleotide binding; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR003593; AAA ATPase.
DR InterPro: IPR011527; ABC membrane 1.
DR InterPro: IPR001140; ABC TM transp.
DR InterPro: IPR003439; ABC transp_like.
DR InterPro: IPR005292; MRP assoc.
DR InterPro: IPR000719; Prot_kinase.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD000006; ABC_transporter; 2.
DR SMART: SM00382; AAA; 2.
DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
DR PROSITE: PS50929; ABC_TMtr; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE: PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE: PS0107; PROTEIN KINASE_ATP; UNKNOWN 1.
KW ATP-binding; Nucleotide-binding; Repeat; Transport.
SQ SEQUENCE 1531 AA; 171792 MW; 746361A71C6158BD CRC64;

Query Match 72.6%; Score 7272; DB 2; Length 1531;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 1402; Conservative 70; Mismatches 59; Indels 0; Gaps 0;

QY 1 MALRGFSADSDSLMDMNTWNTSNPDTFCQFQNTVVMVPCFYLACPFYFLYSRH 60
DB 1 MALRGFSADSDSLPFWMDVSWNTSNPDTFCQFQNTVVMVPCCLMLCFPFYFLYSRH 60
QY 61 DRGIQMTPLNKTALGFLIMVCMADLFYSFMRSGIFLAPVFLVSPTLGITTLA 120
DB 61 DRGIQMTYLNKTALGFLVIMVCMADLFYSFMRSGIKLAPVFLVSPTLGITTLA 120
QY 121 TFLIQLERRKGVSSGIMLTWLVVALVCAIILNSKITALKEDAQVDLPDITPYVFS 180
DB 121 TFLIQLERRKGVSSGIMLTWLVVALVCAIILNSKITALKEDAIDVFDVTFYIYFS 180
QY 181 LLLIQVLVSCSDSPLESETIHDNPPCESASFLSITFMWITGLIVRGYRQPLESD 240
DB 181 LVLILQVLVSCSDSPLESETIHDNPPCESASFLSVTFMWITGLVWGVYRQPLESTD 240
QY 241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVKVYVSSKDPAPCKESSKYDANEVEAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNMKKECAKSKRQKRTIYSSKDPAPCKGSSQYDVANEVEAL 300
QY 301 IVKSPQKEWNSLFRVLYKTGFYFLMSFFKAIHDLMMFSGPOILKLLIKFVNDYAP 360
DB 301 IVKTPQKERESPLFRVLYKTGFYFLMSFFKALHDLMMFAGPEILKLLINFDVNDKAPD 360
QY 361 WQGFYVYVLFVYACIQVLYHOYFHTCFVSGMRKKTAVIGAVYRKALVITNSARKSSTV 420
DB 361 WQGFYVYVLFVYACIQVLYHOYFHTCFVSGMRKKTAVIGAVYRKALVITNSARKSSTV 420
QY 421 GEIYNLMSVDAQRFMDLATYINMIWSAPLOVILALYLMLNGSPSLAGVAVMYLVN 480
DB 421 GEIYNLMSVDAQRFMDLATYINMIWSAPLOVILALYLMLNGSPSLAGVAVMYLVN 480
QY 481 AVMAAMKTKTYOVANHKSKDNRIKLMNEILNGIKVLKLYAMEIAFKDKVLAIROBELKYLK 540
DB 481 AVMAAMKTKTYOVANHKSKDNRIKLMNEILNGIKVLKLYAMEIAFKDKVLAIROBELKYLK 540
QY 541 KSAITSLAVGTFTWCTPFLVALCTPAVYVVTIDENNIILDAQAFVSLAFNLRPFIATLP 600
DB 541 KSAITSLAVGTFTWCTPFLVALCTPAVYVVTIDENNIILDAQAFVSLAFNLRPFIATLP 600

QY 601 MVISSIVQASVSLKRLRIFLSHELEBDSIERRPVXDGGSNTSVNNAFTWARSDEPT 660
DB 601 MVISSIVQASVSLKRLRIFLSHELEBDSIERRPVKDGGGANSITVKNATFTWARSDEPT 660
QY 661 LNSGTSEIPGALVAVVGVQCGKSSLSLALLAMDKVBEHVAIKGSVAIVPQAMQND 720
DB 661 LNSGTSEIPGSLVAVVGVQCGKSSLSLALLAMDKVBEHVAIKGSVAIVPQAMQND 720
QY 721 SLRENIIFGQLEBPYRSYIOACALLPDEIIPSGRTEIGEKVNLSSGQXORVSLAR 780
DB 721 SLRENIIFGQLEBPYRSYIOACALLPDEIIPSGRTEIGEKVNLSSGQXORVSLAR 780
QY 781 AVYSNADIIYLFDDPLSAVDHVGKHPENYIGPKMLKNKTRILVTHSMSTLPQVDYIIV 840
DB 781 AVYCDSDIYLFDDPLSAVDHVGKHPENYIGPKMLKNKTRILVTHSISTLPQVDYIIV 840
QY 841 MSGKISMSQYQELARDDGAFAPFLTYASTBEOBDAENGVTGVSQPGKAMQNGM 900
DB 841 MTGKISMSQYQELARDDGAFAPFLTYASTBEOBDAENGVTGVSQPGKAMQNGM 900
QY 901 LVTDASGKOLORQSSSSYSGDISRHNSGTABLOKAEAKKEFTWKLMEADKAOTGVYL 960
DB 901 LVTDVAGKOLORQSSSSYSGDVSRRHTSTABLOKAGPKNEADKLVADKATQGVYL 960
QY 961 SVYWDYKAIQLFISFLSIFLMCNHVSALASNYMSLMTDDPIVNGTOEHTKVLRSVYG 1020
DB 961 SVYWDYKAIQLFISFLSIFLMCNHVSALASNYMSLMTDDPIVNGTOEHTKVLRSVYG 1020
QY 1021 ALGISQIATVFGSMANVSIGIFASRLHVDLLQNVLRSPSPFERPSGLVNRFSKEL 1080
DB 1021 ALGISQIATVFGSMANVSIGIFASRLHVDLLQNVLRSPSPFERPSGLVNRFSKEL 1080
QY 1081 DTVDMSIPEYIKMFGSLFNVIGACIYLLATPIAIIIPPLGIYFFVQGFYASSROL 1140
DB 1081 DTVDMSIPEYIKMFGSLFNVIGACIYLLATPIAIIIPPLGIYFFVQGFYASSROL 1140
QY 1141 KRLESVSRSPYSHFNELLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
DB 1141 KRLESVSRSPYSHFNELLGVSVIRAFEEQERFIHQSDLKVDENOKAYPSIVANRWLA 1200
QY 1201 VRLCEVNCIYLPALAVISRHSLSAGVGLSVSYSLQVTTYLMVVRMSSEMETIVA 1260
DB 1201 VRLCEVNCIYLPALAVISRHSLSAGVGLSVSYSLQVTTYLMVVRMSSEMETIVA 1260
QY 1261 VERLKEYSETEKEAPMOIOETAPPSWPQYGRVAFRYCARYREDLPVLRHINVTNGG 1320
DB 1261 VERLKEYSETEKEAPMOIOETAPPSWPQYGRVAFRYCARYREDLPVLRHINVTNGG 1320
QY 1321 EKVGIVRTGAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVL 1380
DB 1321 EKVGIVRTGAGKSLTLGLFRINESAGEIIIDGINIAKIGLHDLRFKTIIPQDPVL 1380
QY 1381 SGIIRANLDPFSQYSDBEVWTSLELAHKDFVSAAPKLDHECAEGGENSVGQROLYCL 1440
DB 1381 SGIIRANLDPFSQYSDBEVWTSLELAHKDFVSAAPKLDHECAEGGENSVGQROLYCL 1440
QY 1441 ARALLRKTILVUDEAFAVDLERDDLQSTIRQFEDCTVLTARHLNTIMDTRYIVL 1500
DB 1441 ARALLRKTILVUDEAFAVDLERDDLQSTIRQFEDCTVLTARHLNTIMDTRYIVL 1500
QY 1501 DKGEIOEYGAPSDDLQORGLFYSMADAGLV 1531
DB 1501 DKGEIRREGQPSDDLQORGLFYSMADAGLV 1531

RESULT 8
059G19 HUMAN PRT; 1439 AA.
ID 059G19 HUMAN PRELIMINARY;
AC 059G19;
DT 10-MAY-2005 (Tremblrel. 30, Created)
DT 10-MAY-2005 (Tremblrel. 30, Last sequence update)
DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)

DE ATP-binding cassette, sub-family C, member 1 isoform 1 variant
(Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
NCBI_TaxID=9606;
RN
RP NUCLEOTIDE SEQUENCE.
RC
RA TISSUE=Brain.
RA Tctoki Y., Joyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title."
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL, AB209120, BAD92357.1; -, mRNA.
DR Ensembl, ENSG00000103222; Homo sapiens.
DR GO: GO:0016021; C:Integral to membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0016887; F:ATPase activity; IEA.
DR GO: GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:0000166; F:nucleotide binding; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR011527; ABC membrane 1.
DR InterPro: IPR001140; ABC_TM_transp.
DR InterPro: IPR003439; ABC_TM_transp.
DR InterPro: IPR005292; MRP_assoc.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR SMART, SM00382; AAA; 2.
DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
DR PROSITE: PS00929; ABC_TM1; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE: PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE: PS0107; PROTEIN KINASE_ATP; UNKNOWN_1.
KM ATP-binding; Nucleotide-binding.
FT
SQ SEQUENCE 1439 AA; 160439 MW; 090D881ADBFB3077 CRC64;
Query Match 72.3%; Score 7242; DB 2; Length 1439;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1428; Conservative 0; Mismatches 1; Indels 10; Gaps 1;
QY 103 APFVLSPTLLGTTTLATFLIQLEKRGVQSSGIMLTFWLVALVCAALIRSKIMTALK 162
DB 1 APFVLSPTLLGTTTLATFLIQLEKRGVQSSGIMLTFWLVALVCAALIRSKIMTALK 60
QY 163 EDQVDFRDIITFVYVSSLLIQVLVSCFSDRSPLESETIHDPPCESSASPLSRITFW 222
DB 61 EDQVDFRDIITFVYVSSLLIQVLVSCFSDRSPLESETIHDPPCESSASPLSRITFW 120
QY 223 WITGLVYRGYROPLEGGDLMSLNKEDTSEQVVPVLVQNMKECAKTRKQPVKYVSSKDP 282
DB 121 WITGLVYRGYROPLEGGDLMSLNKEDTSEQVVPVLVQNMKECAKTRKQPVKYVSSKDP 180
QY 283 ACPKESKQVDAVEEALIVKSPQKWNPSLFVLYTTFGPYFLMSFFFAIHDLMWFSG 342
DB 181 ACPKESKQVDAVEEALIVKSPQKWNPSLFVLYTTFGPYFLMSFFFAIHDLMWFSG 240
QY 343 POLIKLILKFNVTKAPDMQGYFTVLVPTACQLTIVLHOYFHCVSGMRIRKAVIGA 402
DB 241 POLIKLILKFNVTKAPDMQGYFTVLVPTACQLTIVLHOYFHCVSGMRIRKAVIGA 300
QY 403 VYRKALVITNSARKSSITVGEIVNLMSVDAQRFMDLATYINMISAPQLVILALYLLMLN 462
DB 301 VYRKALVITNSARKSSITVGEIVNLMSVDAQRFMDLATYINMISAPQLVILALYLLMLN 360
QY 463 GBSVLGAVAVMLVMPNAAVMAMKTKTYQVAHMSKKNRIKLANEILNGIKVLAYAMEL 522
DB 361 GBSVLGAVAVMLVMPNAAVMAMKTKTYQVAHMSKKNRIKLANEILNGIKVLAYAMEL 420

QY 523 AFKDKVLAIROBELKVLKKSAYLSAVGTFTWVCTPFLVALCTFAVYVYIDENNILDAQTA 582
DB 421 AFKDKVLAIROBELKVLKKSAYLSAVGTFTWVCTPFLVALCTFAVYVYIDENNILDAQTA 480
QY 583 FVSLALFNILRFPNLILPMVLISSIVQASVSLKRLIFLSHEELPDSEIRRPVKDGGCTN 642
DB 481 FVSLALFNILRFPNLILPMVLISSIVQASVSLKRLIFLSHEELPDSEIRRPVKDGGCTN 540
QY 643 SITVRNATFTWASDPTLNGITPSTIPBGLVAVVGVGGKSSLSALLAEMDKVGHV 702
DB 541 SITVRNATFTWASDPTLNGITPSTIPBGLVAVVGVGGKSSLSALLAEMDKVGHV 600
QY 703 AIKGSVAVYVQOAMIQNDSLRNENILFGCOLEPPYRSVIOACALLPDLILPSGDRFEIG 762
DB 601 AIKGSVAVYVQOAMIQNDSLRNENILFGCOLEPPYRSVIOACALLPDLILPSGDRFEIG 660
QY 763 EKGVLISGGQKQVRSARAVYSNADYILFPDPLSAVDAAVGHKII FENVIGPKMLKNKTR 822
DB 661 EKGVLISGGQKQVRSARAVYSNADYILFPDPLSAVDAAVGHKII FENVIGPKMLKNKTR 720
QY 823 ILVTHSMSTYLPQVDVITVMSGGKISMGSTYQELLARDGAFELRTYASTEOEDAEEN- 881
DB 721 ILVTHSMSTYLPQVDVITVMSGGKISMGSTYQELLARDGAFELRTYASTEOEDAEEN- 780
QY 882 -----GVTVSGPGKAKQKMGMLVTDGAKGOLQSSSSSGDISRHNSTA 932
DB 781 STWDEBEAGVTVSGPGKAKQKMGMLVTDGAKGOLQSSSSSGDISRHNSTA 840
QY 933 ELQKAEKKEETKWLMEADAKQGVYLSYVMDYKRAIGLIFSLISFLFMCHVSLAS 992
DB 841 ELQKAEKKEETKWLMEADAKQGVYLSYVMDYKRAIGLIFSLISFLFMCHVSLAS 900
QY 993 NYWLSMTDDPIYNGIOEHTKRLSYGALGISQGIANVFYSMAVSGIILASRCHVDL 1052
DB 901 NYWLSMTDDPIYNGIOEHTKRLSYGALGISQGIANVFYSMAVSGIILASRCHVDL 960
QY 1053 LHSILSPMSFERRTSGNLVNRFSKELDVDSMIPEVIMFMSGLFNVGACIVILLAT 1112
DB 961 LHSILSPMSFERRTSGNLVNRFSKELDVDSMIPEVIMFMSGLFNVGACIVILLAT 1020
QY 1113 PIAAIIIPPLGLIYFFQRFYVASRQLKLESVSRSPPVSHFNETLLAGSVIRAFEEQE 1172
DB 1021 PIAAIIIPPLGLIYFFQRFYVASRQLKLESVSRSPPVSHFNETLLAGSVIRAFEEQE 1080
QY 1173 RFIHQSDLKVDENQKAYPSIVANRWLAVERECVGCNCTVPAALFAVISRHSLSAGLVGL 1232
DB 1081 RFIHQSDLKVDENQKAYPSIVANRWLAVERECVGCNCTVPAALFAVISRHSLSAGLVGL 1140
QY 1233 SVYSLSQVTTYLNMVLRMSSEMETNIVAVERLKEYSETEKEAPMOIQETAPSPSPQVGR 1292
DB 1141 SVYSLSQVTTYLNMVLRMSSEMETNIVAVERLKEYSETEKEAPMOIQETAPSPSPQVGR 1200
QY 1293 VEFRRNYCLARYREDIDPVLRIHNTVINGEKVGIIVGRTGAKSSLTGLFRLINSABEII 1352
DB 1201 VEFRRNYCLARYREDIDPVLRIHNTVINGEKVGIIVGRTGAKSSLTGLFRLINSABEII 1260
QY 1353 IDGINIAKIGLHDLRKKITITIPQDPVLFGSGLRPMNLDPSSQYDEEWTSLBLAHLKDFV 1412
DB 1261 IDGINIAKIGLHDLRKKITITIPQDPVLFGSGLRPMNLDPSSQYDEEWTSLBLAHLKDFV 1320
QY 1413 SALPDLKDHCAEGENLSVGRQVLCIARALLRKTIIVLDATAVVDLETDDLQSTI 1472
DB 1321 SALPDLKDHCAEGENLSVGRQVLCIARALLRKTIIVLDATAVVDLETDDLQSTI 1380
QY 1473 RTQFEDCTVLTIAHRLNTIMDYRVVLVDKGEIQEAGPBDLLQOQGLFSPMAKDAGLV 1531
DB 1381 RTQFEDCTVLTIAHRLNTIMDYRVVLVDKGEIQEAGPBDLLQOQGLFSPMAKDAGLV 1439

RESULT 9
Q8HX05_BOVIN
ID Q8HX05_BOVIN PRELIMINARY; PRT; 1530 AA.
AC Q8HX05;

DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Multidrug resistance protein 1.
GN Name=MRP1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22062955; PubMed=12067707; DOI=10.1016/S0014-5793(02)02816-8;
RA Teguchi Y., Saeki K., Komano T.;
RT "Functional analysis of MRP1 cloned from bovine."
RL FBS Lett. 521:211-213(2002).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL: AB082124; BAC15550.1; -; mRNA.
DR HSSP: P08716; IMT0.
DR GO: GO:0016021; C:Integral to membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0016887; F:ATPase activity; IEA.
DR GO: GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:0000166; F:nucleotide binding; IEA.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR011527; ABC membrane 1.
DR InterPro: IPR001140; ABC_TM_transp.
DR InterPro: IPR003439; ABC_transp_like.
DR InterPro: IPR005292; MRP_assoc.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD000006; ABC_transporter; 2.
DR SMART: SMO0382; AAA; 2.
DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
DR PROSITE: PS0929; ABC_TM1F; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE: PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1530 AA; 171666 MW; AAE4F92ED7832703 CRC64;

Query Match 72.2%; Score 7230.5; DB 2; Length 1530;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 1388; Conservative 83; Mismatches 59; Indels 1; Gaps 1;

QY 1 MALSGFCSADGSDLLMDNMTNTNSPDFTKCFQNTVLVWVPCFTLACFPFELYLSRH 60
DB 1 MALDPCFSDGSDLFWEWNVMTNTNSPDFTKCFQNTVLVWVPCFTLACFPFELYLSH 60
QY 61 DRGYIOMTPLNKTALGFLIMIVCMADLFYSFWERSRGIFLAPVFLVSPILLGITLLA 120
DB 61 DRGYIOMTHLNKATLGFLLMIVCMADLFYSFWERSRGKTLAPVFLVSPILLGITLLA 120
QY 121 TFLIQLERRKVOSSGIMLTFWLVALVCAALINSKIMTALKEDAVDLPDITFYVFS 180
DB 121 TFLIQLERRRVOSSGIMLTFWLVALVCAALILSKIMTALKEDARVDVDFVIFYVFS 180
QY 181 LLLIQLVIVSCFSDSPFSETIHDPNCPSSASFLSITFMWITGLVLRGROPLEGSD 240
DB 181 LVLILQLVIVSCFSDSPFSETIHDPNCPSSASFLSITFMWITGMVVOGRQPLEGSD 240
QY 241 LMSLNKEDTSEQVVPVLVKNMKKECAKTRKQPVKVYVSSKDPAPCKESSKYDANEVAL 300
DB 241 LMSLNKEDTSEQVVPVLVKNMKKECAKSRKQPVKVYSSKDPARKGSSKYDVANEVAL 300
QY 301 IVKSPQKEMNSLFPKLVYKTFGYPFLMSFFPKAHLDMFSGPOLIKLLIKFVNDTKAP 360
DB 301 IVKCPQKRDSDSLFVLYKTFGYPFLMSFLFKAVHDLMMFAGPBITLKLINFVNDKAP 360
QY 361 WQGFYTYLLFVTAQCLQTLVHLQVPHILCFVSGMRKTAIVIGAVYKKAIVITNSARKSTV 420
DB 361 WQGFYTYLLFISACLQTLVHLQVPHILCFVSGMRKTAIVIGAVYKKAIVITNAARKSTV 420

QY 421 GEIVNIMSVDAORFMDIATYINNIWSAPLOVILLALYLLMNLGPSVLGAVAVMLNVPV 480
DB 421 GEIVNIMSVDAORFMDIATYINNIWSAPLOVILLALYLLMNLGPSVLGAVAVMLNVPV 480
QY 481 ANWAMKTKYOVAMHMSKDNRIKLMEIILNGIVLTKUYAMELAPKQVLAIRQELVVK 540
DB 481 ANWAMKTKYOVAMHMSKDNRIKLMEIILNGIVLTKUYAMELAPKQVLAIRQELVVK 540
QY 541 KSAVLSAVGTFTWCCEFLVALCTFAVYVITDENNIIDAOATFASLAFNILRFLNILP 600
DB 541 KSAVLSAVGTFTWCCEFLVALCTFAVYVITDENNIIDAOATFASLAFNILRFLNILP 600
QY 601 MVISSIVQASVSLKRLRILFISHEELEPDSIERRPVKDGGTNSITVNAFTWARSDDPT 660
DB 601 MVISSIVQASVSLKRLRILFISHELEPDSIQRRPPIKQAGATNSITVNAFTWARSDDPT 660
QY 661 LNGITFSEPGALVAVVGQCGCKSSILSLALMEVDYBEHVAMIKGSVAVVPQAMIQND 720
DB 661 LHGITFSEPGSLVAVVGQCGCKSSILSLALMEVDYBEHVAMIKGSVAVVPQAMIQNT 720
QY 721 SLRENIILFGQLEBPYRSVYIQAALLPDLIELIPSGDRTIEGKGVNLSGGQKQSVSLAR 780
DB 721 SLRENIILFGQLOERYRYKAVVEACALLPDLIELIPSGDRTIEGKGVNLSGGQKQSVSLAR 780
QY 781 AVYSNADYILPDPPLSAVDAAHVGHKIFENYIGPKMLAKNTRILVTHSMSTLPQVDYIIV 840
DB 781 AVYCDSDVYLLDPLSAVDAAHVGHKIFENYIGPKMLAKNTRILVTHAISLTPQMDYIIV 840
QY 841 MSGGKISEMSYQELLARDAPAEFLRTASTQEDODAEKNGVGVSGPGEAQOMNGM 900
DB 841 MSGGKISEMSYQELLARDAPAEFLRTASTQEDODAEKNGVGVSGPGEAQOMNGM 900
QY 901 LVTSACKQOROLSSSSSYSGDISRRHNSYAELOKAEAKKEETWKLMEADKAQTQVKL 960
DB 901 LVYDTACKQOROLSSSSSYSRDVSCHHTSTAEARK-PGTEETWKLVEADKAQTQVKL 959
QY 961 SVYDWYKAIQLFISFLIFPMCNHVSALASNYLSLMTDDPVLVNGTQHTKRLSVYG 1020
DB 960 SVYDWYKAIQLFISFLIFLPCNHVASLVSNYLSLMTDDPVLVNGTQHTQRLSVYG 1019
QY 1021 ALGISQGIAPFGYSMANSTIGIILASRCLHYDLHSILRSFMSPEETPSGTVNRFSKL 1080
DB 1020 ALGISQGIATFGYSAVASIGIFASRLHLDLHNVRSPISPEKTPSGVLNRFSEKL 1079
QY 1081 DTVDSMIPVYIKMFGSLFNVYIGACIYLLATPIAIIIPPLGIYFFVQRFYASSROL 1140
DB 1080 DTVDSMIPVYIKMFGSLFNVYIGACIITLATPIAAVYIIPPLGIYFFVQRFYASSROL 1139
QY 1141 KRLESVSRSPVYSHFNETLLGVSVYIRAPEEQERFIHQSDLKVDENQKAYYPSIVANFWLA 1200
DB 1140 KRLESVSRSPVYSHFNETLLGVSVYIRAPEEQERFIHQSDLKVDENQKAYYPSIVANFWLA 1199
QY 1201 VRLCEGNCIVLPALFAVISRHSLSAGLVGLSVSYLQVTTYVNLVRRSSSEMETNIVA 1260
DB 1200 VRLCEVNCIVLPALFAVISRHSLSAGLVGLSVSYLQVTTYVNLVRRSSSEMETNIVA 1259
QY 1261 VERLKESETEKEAPMOIQETAPPSWPOYGRVFRNYCLRYREDLDFVLRIHIVTINGG 1320
DB 1260 VERLKESETEKEAPMOIQMAAPPKQMPPOYGRVFRNYCLRYREDLDFVLKHIIVTINGG 1319
QY 1321 EKVGIVRTGAGKSSLTGLFRINESAGEEIIIDGINIAKIGLHDLRPFKTTIIPQDPLF 1380
DB 1320 EKVGIVRTGAGKSSLTGLFRIKESAGEEIIIDGINIAKIGLHDLRPFKTTIIPQDPLF 1379
QY 1381 SSGSRNMLDPPSQSDEDEWTSLELAHLKDFVSLPKLDBHECAGEGNSLVGQROLVCL 1440
DB 1380 SSGSRNMLDPPSQSDEDEWTSLELAHLKGVVSLPKLDBHECAGEGNSLVGQROLVCL 1439
QY 1441 ABALLRKTILVLEAATAVVDLETDDLIQSTIRFOPEDCYLLTAHRLNTIMDTTRVIVL 1500
DB 1440 ABALLRKTILVLEAATAVVDLETDDLIQSTIRFOPEDCYLLTAHRLNTIMDTTRVIVL 1499
QY 1501 DKGEIOBYGAFSDLLQORGLFYMAKDAGLV 1531

QY 977 LSIPLFMCHNSALASNLWLSMTWDDPYNGTQHTKRLSYGALGISOIAVFGYSMA 1036
 DB 846 LSIPLFMCHNSALASNLWLSMTWDDPYNGTQHTKRLSYGALGISOIAVFGYSMA 905
 QY 1037 VSIGGILASRCALVDLHLSILASPMSPFERTSGMLVNRFSKELTVDSMIPEVKMFNG 1096
 DB 906 VSIGGILASRCALVDLHLSILASPMSPFERTSGMLVNRFSKELTVDSMIPEVKMFNG 965
 QY 1097 SLFNIVGACIVILLATPTIAAIIIPPLGLIFFVQRFYVASSHQRLKLESVSPYVSHN 1156
 DB 966 SLFNIVGACIVILLATPTIAAIIIPPLGLIFFVQRFYVASSHQRLKLESVSPYVSHN 1025
 QY 1157 ETLGLSVITRAPEEERFIHQSDLVKVDENQKAYYSISVANRWLAVRLCEVGCIVLPAL 1216
 DB 1026 ETLGLSVITRAPEEERFIHQSDLVKVDENQKAYYSISVANRWLAVRLCEVGCIVLPAL 1085
 QY 1217 FAVISRHSLSAGLVGSLVSYSLQVTTYLNMVLRMSSEMTNVAVERLKEYSETEKAPW 1276
 DB 1086 FAVISRHSLSAGLVGSLVSYSLQVTTYLNMVLRMSSEMTNVAVERLKEYSETEKAPW 1145
 QY 1277 QIQETAPPSWMPQVGRVEFRNYCLRYRBDLDFVLRHINVTINGEKVGIWRTGAGKSSL 1336
 DB 1146 QIQETAPPSWMPQVGRVEFRNYCLRYRBDLDFVLRHINVTINGEKVGIWRTGAGKSSL 1205
 QY 1337 TLGLFRINESAGEIIIDGINIAKIGLHDLRKITITIPDPPLFSGSLRMLDPSQYSD 1396
 DB 1206 TLGLFRINESAGEIIIDGINIAKIGLHDLRKITITIPDPPLFSGSLRMLDPSQYSD 1265
 QY 1397 EEWVTSLELAHKDVPVSLPDLDBECAGGNNLSVGOQLVCLARALLRKTIIVLDEA 1456
 DB 1266 EEWVTSLELAHKDVPVSLPDLDBECAGGNNLSVGOQLVCLARALLRKTIIVLDEA 1325
 QY 1457 TAAVLDLTDLLIQSTIRTOFEDCTVLIARLNTIMDYRVILVLDKGEIOEGYASDILQ 1516
 DB 1326 TAAVLDLTDLLIQSTIRTOFEDCTVLIARLNTIMDYRVILVLDKGEIOEGYASDILQ 1385
 QY 1517 QRGLEYMAKXAGLV 1531
 DB 1386 QRGLEYMAKXAGLV 1400

RESULT 11
 MRPL_MOUSE STANDARD; PRT; 1528 AA.
 AC 035379;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Multidrug resistance-associated protein 1 (ATP-binding cassette sub-family C member 1).
 GN Name=Abcc1; Synonyms=Abcc1a, Abcc1b, Mdrap, Mrp;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridea; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Muscle;
 RC MEDLINE=96251691; PubMed=8649356;
 RA Stride B.D., Valdimarsson G., Gerlach J.H., Wilson G.M., Cole S.P.C., Deeley R.G.;
 RT "Structure and expression of the messenger RNA encoding the murine multidrug resistance protein, an ATP-binding cassette transporter.",
 RL Mol. Pharmacol. 49:962-971(1996).
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RP STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S., Nikaide I., Otsu N., Saito R., Suzuki H., Yamana K.I., Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T., Baldarelli R., Hall D.P., Bult C., Hume D.A., Quackenbush J.,

RA Schriml L.M., Kanapin A., Mateuda H., Betalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S., Dalla E., Dregant T.A., Fletcher C.F., Forrest A., Frazer K.S., Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J., Grimmond S., Guerinich S., Hirokawa N., Jackson I.J., Jarvis E.D., Kani A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L., Konagaya A., Kurochkin I.V., Lee Y., Lehndt B., Lyons P.A., Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H., Nagashima T., Numata K., Okido T., Pavan W.J., Perlea G., Pesole G., Petrovsky N., Pillai R., Pontus J.U., Qi D., Ramchandran S., Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M., Sandelin A., Schneider C., Sempile C.A., Secou M., Shimada K., Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M., Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C., Wilming L.G., Wysshaw-Bois A., Yanagisawa M., Yang I., Yang L., Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carrinci P., Hayatsu N., Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aizawa K., Arikawa T., Fukuda S., Hara A., Hashizume W., Imocant K., Ishii Y., Itoh M., Kagawa I., Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J., Birney E., Hayashizaki Y.;
 RA "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs";
 RT Nature 420:563-573(2002).
 RL
 CC -1- FUNCTION: May participate directly in the active transport of drugs into subcellular organelles or influence drug distribution indirectly (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 CC -1- SIMILARITY: Belongs to the ABC transporter family. MRP subfamily.
 CC -1- SIMILARITY: Contains 2 ABC transmembrane type-1 domains.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
 CC -----
 DR EMBL: AF022908; AAC80938.1; -; mRNA.
 DR EMBL: AK029876; BAC26654.1; -; mRNA.
 DR HSPB: P08716; MTO.
 DR Ensembl: ENSMUSG00000023088; Mus musculus.
 DR MGI: MGI:102676; Abcc1.
 DR GO: GO:0005887; C:integral to plasma membrane; IDA.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR011527; ABC_membrane_1.
 DR InterPro: IPR001140; ABC_TM_transp.
 DR InterPro: IPR003439; ABC_tranp_1like.
 DR InterPro: IPR005292; MRP_assoc.
 DR Pfam: PF00664; ABC_membrane; 2.
 DR Pfam: PF00005; ABC_tranp; 2.
 DR ProDom: PD000006; ABC_transporter; 2.
 DR SMART: SM00382; AAA; 2.
 DR TIGRFAMs: TIGR00957; MRP_assoc_pro; 1.
 DR PROSITE: PS50929; ABC_TM1F; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
 DR PROSITE: PS50893; ABC_TRANSPORTER_2; 2.
 KM ATP-binding; Glycoprotein; Nucleotide-binding; Repeat; Transmembrane; transport.
 KM
 FT TOPO_DOM 1 33 Extracellular (By similarity).
 FT TRANSMEM 34 54 1 (By similarity).
 FT TOPO_DOM 55 74 Cytoplasmic (By similarity).
 FT TRANSMEM 75 95 2 (By similarity).
 FT TOPO_DOM 96 100 Extracellular (By similarity).
 FT TRANSMEM 101 121 3 (By similarity).
 FT TOPO_DOM 122 133 Cytoplasmic (By similarity).
 FT TRANSMEM 134 154 4 (By similarity).
 FT TOPO_DOM 155 172 Extracellular (By similarity).
 FT TRANSMEM 173 193 5 (By similarity).
 FT TOPO_DOM 194 317 Cytoplasmic (By similarity).
 FT TRANSMEM 318 338 6 (By similarity).
 FT TOPO_DOM 339 364 Extracellular (By similarity).

FT TRANSMEM 365 385 7 (By similarity).
FT TOPO DOM 386 441 Cytoplasmic (By similarity).
FT TRANSMEM 442 462 8 (By similarity).
FT TOPO DOM 463 465 Extracellular (By similarity).
FT TRANSMEM 466 486 9 (By similarity).
FT TOPO DOM 487 548 Cytoplasmic (By similarity).
FT TRANSMEM 549 569 10 (By similarity).
FT TOPO DOM 570 591 Extracellular (By similarity).
FT TRANSMEM 592 612 11 (By similarity).
FT TOPO DOM 613 663 Cytoplasmic (By similarity).
FT TRANSMEM 664 984 12 (By similarity).
FT TOPO DOM 985 1022 Extracellular (By similarity).
FT TRANSMEM 1023 1043 13 (By similarity).
FT TOPO DOM 1044 1086 Cytoplasmic (By similarity).
FT TRANSMEM 1087 1107 14 (By similarity).
FT TOPO DOM 1108 1108 Extracellular (By similarity).
FT TRANSMEM 1109 1129 15 (By similarity).
FT TOPO DOM 1130 1200 Cytoplasmic (By similarity).
FT TRANSMEM 1201 1221 16 (By similarity).
FT TOPO DOM 1222 1223 Extracellular (By similarity).
FT TRANSMEM 1224 1244 17 (By similarity).
FT TOPO DOM 1245 1528 Cytoplasmic (By similarity).
FT DOMAIN 326 609 ABC transmembrane type-1.1.
FT DOMAIN 644 868 ABC transporter 1.
FT DOMAIN 971 1253 ABC transmembrane type-1.2.
FT NP_BIND 1290 1524 ABC transporter 2.
FT NP_BIND 678 685 ATP 1 (Potential).
FT NP_BIND 1324 1331 ATP 2 (Potential).
FT CARBOHYD 19 19 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1003 1003 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 1528 AA; 171185 MW; 68PD1367D61DBB CRC64;

Query Match 69.9%; Score 7002.5; DB 1; Length 1528;
Best Local Similarity 88.0%; Pred. No. 0;
Matches 1349; Conservative 102; Mismatches 75; Indels 7; Gaps 5;

QY 1 MALRGFSGADGSDLMNMTWMTSNPDFTGCFONTLVWPCGYLWACPFYLYSRH 60
DB 1 MALRSPGADGSDLMNMTWMTSNPDFTGCFONTLVWPCGYLWACPFYLYSRH 60
QY 61 DRGVIQMTPLNKTALGFLIMVCMADLFYSFMRSGRGLFAPLVFVSPFLIGITLLA 120
DB 61 DRGVIQMTPLNKTALGFLIMVCMADLFYSFMRSGRGLFAPLVFVSPFLIGITLLA 120
QY 121 TPLIQLERRKGVSSGIMLTFWLVVCAALIRSKIMTAKEDAOVDLFDITFYVFS 180
DB 121 TPLIQLERRKGVSSGIMLTFWLVVCAALIRSKIMTAKEDAOVDLFDITFYVFS 180
QY 181 LLIILQLVSCGSDSPLESETIHDNCPRESSASFLRITFMWITGLIVGRQPLSGSD 240
DB 181 LLIILQLVSCGSDSPLESETIHDNCPRESSASFLRITFMWITGLIVGRQPLSGSD 240
QY 241 LMSLNKEDTSBOVAVLVKMKKECAKTRKOPVKKVYS-SKDPAPKRESSKVDANEVEA 299
DB 241 LMSLNKEDTSBOVAVLVKMKKECAKTRKOPVKKVYS-SKDPAPKRESSKVDANEVEA 299
QY 300 LIYKSPKEMNPLSKFLVYKTFPGPYFLMSEFFKAHDLMMFSGBOILKLIKFNVDKAP 359
DB 300 LIYKSPKEMNPLSKFLVYKTFPGPYFLMSEFFKAHDLMMFSGBOILKLIKFNVDKAP 359
QY 360 DMQGYFTVLLFVACQTLVHGYPHICFVSGKRITAVIGAVYRKALVITNSAKSST 419
DB 360 DMQGYFTVLLFVACQTLVHGYPHICFVSGKRITAVIGAVYRKALVITNSAKSST 419
QY 420 VGEIVNLSYDAQGFMDLATYINMIWSAPLOVILATLYLMLNLPSTVLGAVAVVAVPV 479
DB 420 VGEIVNLSYDAQGFMDLATYINMIWSAPLOVILATLYLMLNLPSTVLGAVAVVAVPV 479
QY 480 NAYMAKTKTYQVAHMSKDNRIKLANEILNGIVLVKLYAWELAFKQVLAIRQELKVL 539
DB 480 NAYMAKTKTYQVAHMSKDNRIKLANEILNGIVLVKLYAWELAFKQVLAIRQELKVL 539
QY 540 KKSAYLSAVGFTFWVCIPFLVALCTPAVYVITIDENNILDAQTAVSLAPRIILRPILNLT 599
DB 540 KKSAYLSAVGFTFWVCIPFLVALCTPAVYVITIDENNILDAQTAVSLAPRIILRPILNLT 599

DB 541 KKSAYLSAVGFTFWVCIPFLVALCTPAVYVITIDENNILDAQTAVSLAPRIILRPILNLT 600
QY 600 PMVSIIVQASVSIKRLRIFLSHELEPDSIERPPVODGGTNSITVRNATFTWASDDP 659
DB 601 PMVSIIVQASVSIKRLRIFLSHELEPDSIERPPVODGGTNSITVRNATFTWASDDP 659
QY 660 TUNGTITISIEGALVAVAVGVCGCKSSLLSALLAEMKVGHVAIKSVAVVPOQATON 719
DB 660 TUNGTITISIEGALVAVAVGVCGCKSSLLSALLAEMKVGHVAIKSVAVVPOQATON 719
QY 720 DSLRENTLFCQLEPPYRSVIOACALLPDLEILPSGDRTEIGEGVNLSSGQVQVSLA 779
DB 720 DSLRENTLFCQLEPPYRSVIOACALLPDLEILPSGDRTEIGEGVNLSSGQVQVSLA 779
QY 780 RAVYSNADIVLFDPLSAVDAAHVAKHLENVIGRGMKKNKTRILLVTHSMSYLPQVDVI 839
DB 780 RAVYSNADIVLFDPLSAVDAAHVAKHLENVIGRGMKKNKTRILLVTHSMSYLPQVDVI 839
QY 840 VMSSGKISEMSYVELLARGAPAEFLRTASTQEQDABENGVTGVSQKKEKQWENG 899
DB 840 VMSSGKISEMSYVELLARGAPAEFLRTASTQEQDABENGVTGVSQKKEKQWENG 899
QY 900 MLVTDGAKLOLOROLSSSSYSYSGDISRHNHSTAELQAKAKKEETWKLMEADKQTCQVK 959
DB 900 MLVTDGAKLOLOROLSSSSYSYSGDISRHNHSTAELQAKAKKEETWKLMEADKQTCQVK 959
QY 959 LSVYVDMYKAIIGLFIPLFIMCNHVSALASNYMLSLWTD-PYVNGQETHKVLASV 1018
DB 959 LSVYVDMYKAIIGLFIPLFIMCNHVSALASNYMLSLWTD-PYVNGQETHKVLASV 1018
QY 1019 YGALGISQGIAYRGYSNAVSIIGLILASRCHVDLHSLILSPMSFEERTSGNLYNRFK 1078
DB 1019 YGALGISQGIAYRGYSNAVSIIGLILASRCHVDLHSLILSPMSFEERTSGNLYNRFK 1078
QY 1078 YGALGISQGIAYRGYSNAVSIIGLILASRCHVDLHSLILSPMSFEERTSGNLYNRFK 1078
DB 1078 YGALGISQGIAYRGYSNAVSIIGLILASRCHVDLHSLILSPMSFEERTSGNLYNRFK 1078
QY 1139 QLRKLESVSRSPVYSHNETLGLSVIRAREEORFIHQSDLKVDENQKAYYPSIVANRW 1198
DB 1139 QLRKLESVSRSPVYSHNETLGLSVIRAREEORFIHQSDLKVDENQKAYYPSIVANRW 1198
QY 1198 QLRKLESVSRSPVYSHNETLGLSVIRAREEORFIHQSDLKVDENQKAYYPSIVANRW 1198
DB 1198 QLRKLESVSRSPVYSHNETLGLSVIRAREEORFIHQSDLKVDENQKAYYPSIVANRW 1198
QY 1259 VAVERLKEVSETEKEAPMOIOETAPSPSQVGVVERFNCCLARYEDLDVLRHINTIN 1318
DB 1259 VAVERLKEVSETEKEAPMOIOETAPSPSQVGVVERFNCCLARYEDLDVLRHINTIN 1318
QY 1318 VAVERLKEVSETEKEAPMOIOETAPSPSQVGVVERFNCCLARYEDLDVLRHINTIN 1318
DB 1318 VAVERLKEVSETEKEAPMOIOETAPSPSQVGVVERFNCCLARYEDLDVLRHINTIN 1318
QY 1379 LFGSGLRMLNDPPSOYSDEEEMWTSLEIAHLKDPVSALPDYLDHECAEGENLVSQROLV 1438
DB 1379 LFGSGLRMLNDPPSOYSDEEEMWTSLEIAHLKDPVSALPDYLDHECAEGENLVSQROLV 1438
QY 1438 LFGSGLRMLNDPPSOYSDEEEMWTSLEIAHLKDPVSALPDYLDHECAEGENLVSQROLV 1438
DB 1438 LFGSGLRMLNDPPSOYSDEEEMWTSLEIAHLKDPVSALPDYLDHECAEGENLVSQROLV 1438
QY 1499 VLDKGEIOEYCAPSDLLQQRGLFYSMAKDAGLV 1531
DB 1499 VLDKGEIOEYCAPSDLLQQRGLFYSMAKDAGLV 1531

RESULT 12
Q810E4 RAT
ID Q810E4 RAT PRELIMINARY; PRT; 1532 AA.
AC Q810E4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TREMBLrel. 24, last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, last annotation update)
DE ATP-binding cassette protein C1.
GN Name=Abcc1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=spleen;
RA Yabuuchi H., Takayanagi S., Ishikawa T.;
RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL: AF487549; AAC05437.1; -; mRNA.
DR HSSP: P08716; IMTO.
DR Ensembl: ENSRNOG00000032748; Rattus norvegicus.
DR RGD: 3112; Abcc1.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0016887; F:ATPase activity; IEA.
DR GO: GO:0042626; F:ATPase activity, coupled to transmembrane m. .; IEA.
DR GO: GO:0000166; F:nucleotide binding; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR011527; ABC membrane 1.
DR InterPro: IPR001440; ABC_TM_transp.
DR InterPro: IPR003439; ABC_transp_like.
DR InterPro: IPR001395; Aldo/ket_red.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR Prodom: PD000006; Abc transporter; 2.
DR SMART: SM00382; AAA; 2.
DR TIGRFam: TIGR00957; MRP_assoc_pro; 1.
DR PROSITE: PS00929; ABC_TM1F; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE: PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE: PS00063; ALDO_KETO_REDUCTASE_3; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1532 AA; 171493 MW; 2E6939F63F5A3F68 CRC64;

Query Match 69.2%; Score 6932.5; DB 2; Length 1532;
Best Local Similarity 87.3%; Pred. No. 0;
Matches 1339; Conservative 99; Mismatches 92; Indels 3; Gaps 3;

QY 1 MALRGFCAGDSGSDPLMDNNTWNTSNPDFTKCFONTVLVWPCFYLMACEPPYFLYLSRH 60
DB 1 MALSFCCSSDSDPLMDNNTWNTSNPDFTKCFONTVLVWPCFYLMACEPPYFLYLSRH 60
QY 61 DRGYQMTPLNKTALGFELMIVCMADLFYSFMRSRGCTPLARPFLVSPILLGTTTLA 120
DB 61 DRGYQMTPLNKTALGFELMIVCMADLFYSFMRSRGCTPLARPFLVSPILLGTTTLA 120
QY 121 TFLDLERRKGVSSGIMLTFWLVALVCALILRSKIMTALKEDAQVDLFRDITYVYFS 180
DB 121 TFLDLERRKGVSSGIMLTFWLVALVCALILRSKIMTALKEDAQVDLFRDITYVYFS 180
QY 121 TFLDLERRKGVSSGIMLTFWLVALVCALILRSKIMTALKEDAQVDLFRDITYVYFS 180
DB 121 TFLDLERRKGVSSGIMLTFWLVALVCALILRSKIMTALKEDAQVDLFRDITYVYFS 180
QY 181 LILQLVLVSCFSDRSPFLFSETIHDNPNCPSSASFLSRITFWITGLIVGRQPLESGD 240
DB 181 LILQLVLVSCFSDRSPFLFSETIHDNPNCPSSASFLSRITFWITGLIVGRQPLESGD 240
QY 181 LILQLVLVSCFSDRSPFLFSETIHDNPNCPSSASFLSRITFWITGLIVGRQPLESGD 240
DB 181 LILQLVLVSCFSDRSPFLFSETIHDNPNCPSSASFLSRITFWITGLIVGRQPLESGD 240
QY 241 LMSLNKERTSEOVVVLVNMKKECAKTRKQPVKYVYS-SKDPAPKESKIDANEEVZA 299
DB 241 LMSLNKERTSEOVVVLVNMKKECAKTRKQPVKYVYS-SKDPAPKESKIDANEEVZA 299
QY 241 LMSLNKERTSEOVVVLVNMKKECAKTRKQPVKYVYS-SKDPAPKESKIDANEEVZA 300
DB 241 LMSLNKERTSEOVVVLVNMKKECAKTRKQPVKYVYS-SKDPAPKESKIDANEEVZA 300
QY 300 LIVKSPKEMNPSLKFVLYKTGPFYFMSFFPKAIDHLMFSGPOLIKLIFVNDTKAP 359
DB 300 LIVKSPKEMNPSLKFVLYKTGPFYFMSFFPKAIDHLMFSGPOLIKLIFVNDTKAP 359
QY 301 LIVKSHKDRDPSLKFVLYKTGPFYFMSFFPKAIDHLMFSGPOLIKLIFVNDTKAP 360
DB 301 LIVKSHKDRDPSLKFVLYKTGPFYFMSFFPKAIDHLMFSGPOLIKLIFVNDTKAP 360
QY 360 DWQGYFYTVLLFVTACLOTLLVLAHQYFHI CFVSGMRLKTAIVAGAVYRKALVITNSARKSST 419
DB 360 DWQGYFYTVLLFVTACLOTLLVLAHQYFHI CFVSGMRLKTAIVAGAVYRKALVITNSARKSST 419

DB 361 DWQGYFYTVLLFVSACLOTLLVLAHQYFHI CFVSGMRLKTAIVAGAVYRKALVITNSARKSST 420
QY 420 VGEIYVNLMSVDAQDFMDLATYINMIWSAPLOVIALYLMNLGSPVLAGAVVWLVMPV 479
DB 421 VGEIYVNLMSVDAQDFMDLATYINMIWSAPLOVIALYLMNLGSPVLAGAVVWLVMPV 480
QY 480 NAWAMKTKTYQVAMKMSKONRIKLMNEILNGIKYLVKLYAMELAFKDKVLAIRQELKVL 539
DB 481 NAWAMKTKTYQVAMKMSKONRIKLMNEILNGIKYLVKLYAMELAFKDKVLAIRQELKVL 540
QY 540 KKSAYLSAVGTFPTWVCPPPLVALCTPAVYVYTIIDENNILDAQTAVASLAFPIILFPNITL 599
DB 541 KKSAYLSAVGTFPTWVCPPPLVALCTPAVYVYTIIDENNILDAQTAVASLAFPIILFPNITL 600
QY 600 PMVTSIYQSVSLKRLRIFLSHELEPDSIERAPVKGSGTNSITYANAFPTARSDP 659
DB 601 PMVTSIYQSVSLKRLRIFLSHELEPDSIERAPVKGSGTNSITYANAFPTARSDP 660
QY 660 TLNGITFSPREGALVAVGVGCGKSSLSALLAEMDKVEGHVAKGSVAVYPOQAWIQN 719
DB 661 TLNGITFSPREGALVAVGVGCGKSSLSALLAEMDKVEGHVAKGSVAVYPOQAWIQN 720
QY 720 DSIARENILFGQLEBPYRYIQAALLPLELIPSGDRMEIGKGVNLSCGKORVSLA 779
DB 721 DSIARENILFGQLEBPYRYIQAALLPLELIPSGDRMEIGKGVNLSCGKORVSLA 780
QY 780 RAYVSNADITYFDDPLSAVDNAGKHIFENYIGPKMKKNTIRILVTHSMYSLPQVDYII 839
DB 781 RAYVSNADITYFDDPLSAVDNAGKHIFENYIGPKMKKNTIRILVTHSMYSLPQVDYII 840
QY 840 VMSGKISEMGSYQELIARDGAPAEPLRTYASTEOEDAEENYGVSGPKKAKOMENG 899
DB 841 VMSGKISEMGSYQELIARDGAPAEPLRTYASTEOEDAEENYGVSGPKKAKOMENG 900
QY 900 MLVYDSAGKQIQRLSSSSSYSGDISRHNHNTAQLQRAEAKKETWKLMEADKQOTGVK 959
DB 901 MLVYDSAGKQIQRLSSSSSYSGDISRHNHNTAQLQRAEAKKETWKLMEADKQOTGVK 960
QY 960 ILVYDAVGKPIQRLSSSSSYSGDISRHNHNTAQLQRAEAKKETWKLMEADKQOTGVK 959
DB 961 ILVYDAVGKPIQRLSSSSSYSGDISRHNHNTAQLQRAEAKKETWKLMEADKQOTGVK 960
QY 960 LSVYVDYKAKGLFISFLIFLPMCNHYSAASANYWLSLMTDD-PYVNGTOEHFKVRLSV 1018
DB 960 LSVYVDYKAKGLFISFLIFLPMCNHYSAASANYWLSLMTDD-PYVNGTOEHFKVRLSV 1019
QY 1019 YGALGISQGLAVFGYSMAVSIIGGILASRCLHVDLLHSILNSPMSFPERTPSGNLVNRFK 1078
DB 1020 YGALGISQGLAVFGYSMAVSIIGGILASRCLHVDLLHSILNSPMSFPERTPSGNLVNRFK 1079
QY 1079 ELDTYDSMTPEVYIKMFQNSLFNVIACVILLAPPIAIIIPPLGLIYFFQRPYVASSR 1138
DB 1080 ELDTYDSMTPEVYIKMFQNSLFNVIACVILLAPPIAIIIPPLGLIYFFQRPYVASSR 1139
QY 1139 QLKRLSEVSRSPVYSHFNETLLGVSVIRAPEROERFIHQSDLKVDENOKAYYPSIVANRW 1198
DB 1140 QLKRLSEVSRSPVYSHFNETLLGVSVIRAPEROERFIHQSDLKVDENOKAYYPSIVANRW 1199
QY 1199 LAVRLCEVNCIVLPAALFAVISRHSLSAGLVGVSYSLOVTTYLNLVMSSEMETNI 1258
DB 1200 LAVRLCEVNCIVLPAALFAVISRHSLSAGLVGVSYSLOVTTYLNLVMSSEMETNI 1259
QY 1259 VAVERLKEYSTEEKAPWQIOETAPPSGWPOYGRVPEFNYCLARREDLDFLRHINVTIN 1318
DB 1260 VAVERLKEYSTEEKAPWQIOETAPPSGWPOYGRVPEFNYCLARREDLDFLRHINVTIN 1319
QY 1319 GGEKVGIVRTGAGKSSLTGLPFINESABEIIIDGINIKIGLHNRPFKTTIIPQDPV 1378
DB 1320 GGEKVGIVRTGAGKSSLTGLPFINESABEIIIDGINIKIGLHNRPFKTTIIPQDPV 1379
QY 1379 LFSGSLRANLDPFSQYSDDEEYWTSLAHLADFYSAIPDKLDHCABEGENLSVGORQV 1438
DB 1380 LFSGSLRANLDPFSQYSDDEEYWTSLAHLADFYSAIPDKLDHCABEGENLSVGORQV 1439
QY 1439 CLARALLKTKILVDEATAVADLETDDLIOSTRTOEDCTVLTIAHRLNTIMDYTVI 1498
DB 1440 CLARALLKTKILVDEATAVADLETDDLIOSTRTOEDCTVLTIAHRLNTIMDYTVI 1499

QY 1319 GGEKVGIVGRTGAGKSSLTGLFRINESAGEBIIIDGINIAKIGHLDLPFKITIIIPQDPV 1378
DB 1320 GGEKVGIVGRTGAGKSSLTGLFRINESAGEBIIIDGINIAKIGHLDLPFKITIIIPQDPV 1379
QY 1379 LFPSSILRNNDLPFGQYSDEEVTWTSLELAHLKDFVSALPDKLDHECAEGENLSVQGRDLV 1438
DB 1380 LFPSSILRNNDLPFGQYSDEEVTWTSLELAHLKDFVSALPDKLDHECAEGENLSVQGRDLV 1439
QY 1439 CLARALLRKTILVDEATAVNDLETDLLIOSTIRTFQEDCTVLTIAHRLNTIMDYTRVI 1498
DB 1440 CLARALLRKTILVDEATAVNDLETDLLIOSTIRTFQEDCTVLTIAHRLNTIMDYTRVI 1499
QY 1499 VLDKGEIIEYCAPSDLLQORGLFYSMAVADGLV 1531
DB 1500 VLDKGEIIEYCAPSDLLQORGLFYSMAVADGLV 1532
RESULT 14
Q810G9 RAT PRELIMINARY; PRT; 1523 AA.
ID 0810G9 RAT PRELIMINARY; PRT; 1523 AA.
AC 0810G9;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE ATP-binding cassette protein C1 variant A.
GN Name=Abcc1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Spleen;
RA Takayanagi S., Ishikawa T.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AY174892; AAO44983.1; -, mRNA.
DR HSSP; P08716; IMT0.
DR RGD; 3112; Abcc1.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR01527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_TM_transp.
DR InterPro; IPR001395; Aldo/ket_red.
DR InterPro; IPR005292; MRP_assoc.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR PRODom; PD000006; ABC_transporter; 2.
DR SMART; SMO0382; AAA; 2.
DR TIGRFAMs; TIGR00957; MRP_assoc_pro; 1.
DR PROSITE; PS00929; ABC_TM1F; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00063; ALDOXETO_REDUCTASE_3; UNKNOWN_1.
KW ATP-binding; Nucleotide-binding; Repeat.
SQ SEQUENCE 1523 AA; 170505 MW; E40337051A1CB96 CRC64;
Query Match 68.8%; Score 6892; DB 2; length 1523;
Best Local Similarity 87.0%; Pred. No. 0;
Matches 1333; Conservative 98; Mismatches 90; Indels 12; Gaps 4;

DB 61 DRGYIOWTHLNKATTAAGFLMIIICWADLFYSFWERSQGMILAVLLVSPLLGIITMLLA 120
QY TELIQERRKGVSSGIMLTFMLVALVCALAIISKMTALKEADQVDFRDIIFYVYS 180
DB 121 TELIQERRKGVSSGIMLTFMLVALVCALAIISKMTALKEADQVDFRDIIFYVYS 180
QY 121 TELIQERRKGVSSGIMLTFMLVALVCALAIISKMTALKEADQVDFRDIIFYVYS 180
QY 181 LLLIQVLVSCFSDRSPFSETIHDNPNCPRESSASFLRITFMWITGILVRGQRPLEGSD 240
DB 181 LVPIQLVILSCFSDRSPFSETIHDNPNCPRESSASFLRITFMWITGIMVQGRPLKSSD 240
QY 241 LMSLNKEDTSQVVPVLYKMKKECATRKQPVVYVS-SKDPQAPRESSVVDANEVEA 299
DB 241 LMSLNKEDTSQVVPVLYKMKKECATRKQPVVYVS-SKDPQAPRESSVVDANEVEA 299
QY 300 LTVSPOKEMNPSSLFKVLYKTFEGYPILMSFFFKAIHDLMMSGPOILKLKFPVNDTKAP 359
DB 301 LTVSPOKEMNPSSLFKVLYKTFEGYPILMSFFFKAIHDLMMSGPOILKLKFPVNDTKAP 359
QY 360 DMQGYFYVLLFVTACIQTLVLAHQYFHLCEVSGMRKITAIVAGAYRKALVITNSARKSST 419
DB 361 DMQGYLVTLALFVSAQCQTALAHQYFHLCEVSGMRKITAIVAGAYRKALVITNSARKSST 420
QY 420 VGEIYVNLMSVDQAQFPMDLATTYNNWNSAPLOVILALYMLNLGSPVLAGAIVAWLWVPV 479
DB 421 VGEIYVNLMSVDQAQFPMDLATTYNNWNSAPLOVILALYMLNLGSPVLAGAIVAWLWVPV 480
QY 480 NAWMAKTKTYQVAMHMSKDKRIKLMBEILNGIKVILKYAMELAFKQVLAIROBELVYL 539
DB 481 NAWMAKTKTYQVAMHMSKDKRIKLMBEILNGIKVILKYAMELAFKQVLAIROBELVYL 540
QY 540 KKSAYLSAVGTFTWCTPFLVALCTFAVYVITIDENNILDAQTAVSLAFNILREPLNIL 599
DB 541 KKSAYLSAVGTFTWCTPFLVALCTFAVYVITIDENNILDAQTAVSLAFNILREPLNIL 600
QY 600 PMVTSISIVQASVSLKRLRIFLSHELEPDSIERPPVKDGGTNSITTRNAFTTAKRSPP 659
DB 601 PMVTSISIVQASVSLKRLRIFLSHELEPDSIERPSIKDGGGMSNITVGNATFTARAREPP 660
QY 660 TLNLTGTSIPGALVAVVGOVGGCKSSLASLLAEMDKVEGHVAKGSVAVPOOAWIQQ 719
DB 661 TLNLTGTSIPGALVAVVGOVGGCKSSLASLLAEMDKVEGHVAKGSVAVPOOAWIQQ 720
QY 720 DSLRENILFGQLEPEPYRYVYQACALLPDLEILPSPGDRTEIGKGVNLSCGQKORVSLA 779
DB 721 DSLRENILFGQLEPEPYRYVYQACALLPDLEILPSPGDRTEIGKGVNLSCGQKORVSLA 780
QY 780 RAVYSNADYIFDDPLSAVDHVGKHIFBNVYIGPKGMKKNKTRILLVTHSMGYLPQVDYII 839
DB 781 RAVYCNSDIYLLDDPLSAVDHVGKHIFEKVVGGMGLKNKTRILLVTHGISYLPQVDYII 840
QY 840 VMSGKTSSEMGSYOELLARQAFAPLEIRTYASTEOBDAEENGVYSGGPKKAKOMENG 899
DB 841 VMSGKTSSEMGSYOELLARQAFAPLEIRTYASTEOBDAEENGVYSGGPKKAKOMENG 900
QY 900 MLVYDSAGKQOROLSSSSSYSGDISRHHNSTAELQKAEKKEBTWKLMEADKAQOTGVYK 959
DB 901 ILVYDAVAKPL-----HSVYTNQOHSSTALQGS-GYKEETWKLMEADKAQOTGVYK 950
QY 960 LSVYWDYKKAIGLPSLSTSLFPMCNDHVSALASNYWLSLWTD-DPIYNGTOEHTKRLSV 1018
DB 961 LSVYWNWYKKAIGLPSLSTSLFPMCNDHVSALASNYWLSLWTD-DPIYNGTOEHTKRLSV 1010
QY 1019 YGALGISGAGVAFGYSMAVSTGGLASCLHVDLLHSTLRSPMSPFRTSPGNNVNRFSK 1078
DB 1011 YGALGIIQGVAVFYSMAVSTGGLFASRRLLHLDLLQNVLRSPMSPFRTSPGNNVNRFSK 1070
QY 1079 ELDPYDSMIPEVIKMFMSLFFNVIGACIVILATPIAAIIPPLGLIYFFQRFVASSR 1138
DB 1071 ELDPYDSMIPEVIKMFMSLFFNVIGACIVILATPIAAIIPPLGLIYFFQRFVASSR 1130
QY 1139 QLRLEBSVRSRPVYSHFNETLLGVSVIRAFEEQERFIHQSDLKVDENQKAYPSIVANRW 1198

D 1131 QLKELSESRSRSPVSHFNETLLGVS VIRAFEEQERFRQSDLKVDENQKAYPSIVANRW 1190
Q 1199 LAVLEECVNCIYVPAALFAVISRHSLSAGLVGSVSLQVTTYTNLVMSSEMENI 1258
D 1191 LAVLEECVNCIYVPAALFAVISRHSLSAGLVGSVSLQVTTYTNLVMSSEMENI 1250
Q 1259 VAVERLKEYSETEKEAPWQIOETAPSSWPQVGRVEFPNYCCARREDLDFLRHINNTIN 1318
D 1251 VAVRLKEYSETEKEASWQIOETAPSPWHSGRVEFPDYCLRRREDLVKHNITIE 1310
Q 1319 GGEKVGIVGRTGAKSSITLGLFRINESAGEEIIIDGINAKIGLHDLRFKITIIPDPV 1378
D 1311 GGEKVGIVGRTGAKSSITLGLFRINESAGEEIIIDGINAKIGLHDLRFKITIIPDPV 1370
Q 1379 LFGSSILMNLDLPSQYSDSEEWMTSLHLNDLPYSALPDKLDHCEAGENLSSVGORLV 1438
D 1371 LFGSSILMNLDLPSQYSDSEEWMTSLHLNDLPYSALPDKLDHCEAGENLSSVGORLV 1430
Q 1439 CLARALLRKTKIYLVDETAADVLETDLLIOSTIRTOPEDCVTLIAHRLNTIMDYPRVI 1498
D 1431 CLARALLRKTKIYLVDETAADVLETDLLIOSTIRTOPEDCVTLIAHRLNTIMDYPRVI 1490
Q 1499 VLDKGEIOEYGAPSDLLQORGLFYSMADGLV 1531
D 1491 VLDKGEIOEYGAPSDLLQORGLFYSMADGLV 1523

RESULT 15
ID OSF364 CHICK PRELIMINARY; PRT; 1525 AA.
AC OSF364;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
OS ORFNames=RCUIMB04.32d20;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezubov Y., Zaim J.,
RA Fiedler P., Kuter S., Blasodatski A., Kostovska D., Kocer M.,
RA Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.,
RT "Full-length cDNAs from chicken bursa lymphocytes to facilitate
RT genefunction analysis";
RL Genome Biol. 6:R6-R6(2005).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AJ851786; CAH65420.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp.
DR InterPro; IPR005292; NRP_assoc.
DR InterPro; IPR00719; Prot_kinase.
DR Pfam; PF00064; ABC_membrane; 2.
DR Pfam; PF00064; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR TIGRPFAM; TIGR00957; NRP_assoc_pro; 1.
DR PROSITE; PS50929; ABC_TMIF; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.

KW ATP-binding; Hypothetical protein; Nucleotide-binding; Repeat;
KW Transport. 1525 AA; 170972 MW; 387245B711B878FD CR664;
SQ SEQUENCE 1525 AA; 170972 MW; 387245B711B878FD CR664;
Query Match 62.9%; Score 6301; DB 2; Length 1525;
Best Local Similarity 77.6%; Pred. No. 0;
Matches 1193; Conservative 176; Mismatches 148; Indels 20; Gaps 7;
Q 1 MALRGFSADSDLDLNMNTVNTNNTSNDPFCFQNTLVVWPFCYLAACFPFYLYLSRH 60
D 1 MGISLCSADASEPFWMTNLMTHTENPDFTQCFQNTLVVWPFCYLAACFPFYLYLSRH 60
Q 61 DRGYQMTPLNKTALGFLMIYCMADLPSFMSRSGPLAIVPLVSPFLTGITTLTA 120
D 61 DRGYQMTPLNKTALGFLMIYCMADLPSFMSRSGPLAIVPLVSPFLTGITTLTA 120
Q 121 TFLIQLRRKQVSSGIMLTFWVALVCAALIRSKIMTALKEBDAQVDLFRDITFYVYS 180
D 121 TFLIQLRRKQVSSGIMLTFWVALVCAALIRSKIMTALKEBDAQVDLFRDITFYVYS 180
Q 121 TFLIQLRRKQVSSGIMLTFWVALVCAALIRSKIMTALKEBDAQVDLFRDITFYVYS 180
D 121 TFLIQLRRKQVSSGIMLTFWVALVCAALIRSKIMTALKEBDAQVDLFRDITFYVYS 180
Q 181 LLLIQLVLSGSDRSLPSETIHDNPNCPSSSASFLSRITFWITGLIVKGYRQPLSGSD 240
D 181 LLLIQLVLSGSDRSLPSETIHDNPNCPSSSASFLSRITFWITGLIVKGYRQPLSGSD 240
Q 181 LLLIQLVLSGSDRSLPSETIHDNPNCPSSSASFLSRITFWITGLIVKGYRQPLSGSD 240
D 181 LLLIQLVLSGSDRSLPSETIHDNPNCPSSSASFLSRITFWITGLIVKGYRQPLSGSD 240
Q 241 LMSLNKEDTSEIYVGLAKRWAKEMAKTKOPLMLLYSK- - - -KOQKSDSNGEYEE 295
D 241 LMSLNKEDTSEIYVGLAKRWAKEMAKTKOPLMLLYSK- - - -KOQKSDSNGEYEE 295
Q 297 VEALIVSPKEMWNPSEFLKYLYKTPGYFLMSFEFFKAIDLMFSPGOILKLIKFVNDT 356
D 297 VEALIVSPKEMWNPSEFLKYLYKTPGYFLMSFEFFKAIDLMFSPGOILKLIKFVNDT 356
Q 296 AEAALIKPSQSSASISLKYLYKTPGYFLMSFEFFKAIDLMFSPGOILKLIKFVNDT 355
D 296 AEAALIKPSQSSASISLKYLYKTPGYFLMSFEFFKAIDLMFSPGOILKLIKFVNDT 355
Q 357 KAPDWQGYFYTVLIFVTAQIQLVLAHQYFHI CPVSGKRIKTAIVAGVYRRALVITNSARK 416
D 357 KAPDWQGYFYTVLIFVTAQIQLVLAHQYFHI CPVSGKRIKTAIVAGVYRRALVITNSARK 415
Q 356 SAPWQGYFYTVLIFVTAQIQLVLAHQYFHI CPVSGKRIKTAIVAGVYRRALVITNSARK 415
D 356 SAPWQGYFYTVLIFVTAQIQLVLAHQYFHI CPVSGKRIKTAIVAGVYRRALVITNSARK 415
Q 417 SSTWGEIYVMSVDAQFQMLATYINNIWAPLOVILALVILMTNLGSPVLAAGVAVWM 476
D 417 SSTWGEIYVMSVDAQFQMLATYINNIWAPLOVILALVILMTNLGSPVLAAGVAVWM 475
Q 416 TSTVGEIYVMSVDAQFQMLATYINNIWAPLOVILALVILMTNLGSPVLAAGVAVWM 475
D 416 TSTVGEIYVMSVDAQFQMLATYINNIWAPLOVILALVILMTNLGSPVLAAGVAVWM 475
Q 477 VPVNAVAMAKTKYTVQVAMHSKDNRIKLMNBIINGIKYLVKLYAMELAFKQVILAROBEL 536
D 477 VPVNAVAMAKTKYTVQVAMHSKDNRIKLMNBIINGIKYLVKLYAMELAFKQVILAROBEL 535
Q 476 VPVNAVAMAKTKYTVQVAMHSKDNRIKLMNBIINGIKYLVKLYAMELAFKQVILAROBEL 535
D 476 VPVNAVAMAKTKYTVQVAMHSKDNRIKLMNBIINGIKYLVKLYAMELAFKQVILAROBEL 535
Q 537 KVLKSAVYAGVFTWVCTPPLVALCTPAVYVYTIDENNILDAQTAFAVSLAFNLRPPL 596
D 537 KVLKSAVYAGVFTWVCTPPLVALCTPAVYVYTIDENNILDAQTAFAVSLAFNLRPPL 595
Q 536 KVLKSAVYAGVFTWVCTPPLVALCTPAVYVYTIDENNILDAQTAFAVSLAFNLRPPL 595
D 536 KVLKSAVYAGVFTWVCTPPLVALCTPAVYVYTIDENNILDAQTAFAVSLAFNLRPPL 595
Q 597 NILPMVTSIVQASVSLKRLRFLSHSELPPDSITERPVVDGGTNSITVRNATFTARS 656
D 597 NILPMVTSIVQASVSLKRLRFLSHSELPPDSITERPVVDGGTNSITVRNATFTARS 655
Q 596 NILPMVTSIVQASVSLKRLRFLSHSELPPDSITERPVVDGGTNSITVRNATFTARS 655
D 596 NILPMVTSIVQASVSLKRLRFLSHSELPPDSITERPVVDGGTNSITVRNATFTARS 655
Q 657 DPPTLNGITFISREGALVAVVGVGCGKSSLSLALALEMPKVEGHVAIKGSVAVVPOAV 716
D 657 DPPTLNGITFISREGALVAVVGVGCGKSSLSLALALEMPKVEGHVAIKGSVAVVPOAV 715
Q 654 DPPTLNGITFISREGALVAVVGVGCGKSSLSLALALEMPKVEGHVAIKGSVAVVPOAV 713
D 654 DPPTLNGITFISREGALVAVVGVGCGKSSLSLALALEMPKVEGHVAIKGSVAVVPOAV 713
Q 717 IONDSIRENLIIFGQGLEEPPYRSVIOACALLPDLLEIPSGDRTEIGEGVNLSSGQKORV 776
D 717 IONDSIRENLIIFGQGLEEPPYRSVIOACALLPDLLEIPSGDRTEIGEGVNLSSGQKORV 775
Q 714 IONATLEDNIIIFGEMNESRYKRVIEACALLPDLLEIPSGDRTEIGEGVNLSSGQKORV 773
D 714 IONATLEDNIIIFGEMNESRYKRVIEACALLPDLLEIPSGDRTEIGEGVNLSSGQKORV 773
Q 777 SLARAVYVSNADILYFDPLSLAVDAHVKHIFENVIGKGMUKKTRILVYHSMKSYLPOVD 836
D 777 SLARAVYVSNADILYFDPLSLAVDAHVKHIFENVIGKGMUKKTRILVYHSMKSYLPOVD 835
Q 774 SLARAVYVSNADILYFDPLSLAVDAHVKHIFENVIGKGMUKKTRILVYHSMKSYLPOVD 833
D 774 SLARAVYVSNADILYFDPLSLAVDAHVKHIFENVIGKGMUKKTRILVYHSMKSYLPOVD 833
Q 837 VTIIVSGKISKEMSVYQELIARDAFAFEFRTYASTQEOBDAEENGVTGVSQP- GKEAKQ 895
D 837 VTIIVSGKISKEMSVYQELIARDAFAFEFRTYASTQEOBDAEENGVTGVSQP- GKEAKQ 895
Q 834 TIIIVMDGEISEMSYQELIARDAFAFEFRTYASTQEOBDAEENGVTGVSQP- GKEAKQ 888
D 834 TIIIVMDGEISEMSYQELIARDAFAFEFRTYASTQEOBDAEENGVTGVSQP- GKEAKQ 888
Q 896 MENGMVYDSAGQOLROLSSSSSYSQDI- -SRHNSTAELOQAEAKKEBTWKLMEADKA 953
D 896 MENGMVYDSAGQOLROLSSSSSYSQDI- -SRHNSTAELOQAEAKKEBTWKLMEADKA 953
Q 889 VENGVLVNDATGKILMRQLSNNSYSETEKSHOSSTAELOQAEAKKEBTWKLMEADKA 947
D 889 VENGVLVNDATGKILMRQLSNNSYSETEKSHOSSTAELOQAEAKKEBTWKLMEADKA 947
Q 954 QTOGVKLAVYVMDVIMKALIGLIFSLIFLPMCHNVSAALASVYMLSLMWDDITVNGTOHTK 1013
D 954 QTOGVKLAVYVMDVIMKALIGLIFSLIFLPMCHNVSAALASVYMLSLMWDDITVNGTOHTK 1013

```

Db      948  KTRGVKATVWEYMAIGLYISFLSVFLMCMNHASLASNYWLSLMTDDPVVNGQYTN 1007
Qy      1014  VRLSVYGALGISOGIAVFGYSMAVSIIGLILASRCHVDLHLSIRSPMSFFERTPSGNLV 1073
Db      1008  VRLGVYGNLIGISOGIAVFGYSMAVSIIGLILASRCHVDLHLSIRSPMSFFERTPSGNLV 1067
Qy      1074  NRSFKELDTVDSMIPEVILKMFWSLFNVIAGACIVILATPIAIIIPPLGLYFFVQRFY 1133
Db      1068  SRFSEIDITIDSTIPPIIKMFWSGFENVIGACTIILATPIAAVVIIPPLGLVYLLVQRFY 1127
Qy      1134  VASSRQLRLSVSRSPVYSHNETLLGVSVIRAFEOERFIHQSDLYDENOKAYPSI 1193
Db      1128  VATSRQLRLSVSRSPVYSHNETLLGVSVIRAFEOERFIHQSDLYDENOKAYPSI 1187
Qy      1194  VANRLAVRLCEVGNLCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNWLVRMSSE 1253
Db      1188  VANRLAVRLCEVGNLCIVLFAALFAVISRHSLSAGLVGLSVSYSLQVTTYLNWLVRMTSD 1247
Qy      1254  METNIVAVERLKEYSETEKEAPMOIQETAPPSMPQVGRVREFRNYCLARYREDLDFVLRHI 1313
Db      1248  LETNIVAVERYVEYAEEMEKEAEMWSIETAPASTWFOEGKVEFRGFLRYREDLDLVLRNI 1307
Qy      1314  NVTINGEKVGVGRTGAGKSLTGLPRINSAEGEIIIDGINIAKTGLHDLRPKITII 1373
Db      1308  NVTINGEKVGVGRTGAGKSLTGLPRINSAEGEIIIDGINIAKTGLHDLRPKITII 1367
Qy      1374  PODPVLFGSLRMLNDPFSOYSDEEVMTSLELAHLKDFVSALPDKLDHECAEGENLSVG 1433
Db      1368  PODPVLFGSLRMLNDPFDQHSDEDIWRSLELAHLKDFVSALPDKLDHECSBGENLSVG 1427
Qy      1434  QROLVCLARALIRKTKIILVLEATAVADLETDDLIOSTIRTOFEDCTVLTIAHRLNTIMD 1493
Db      1428  QROLVCLARALIRKTKIILVLEATAVADLETDDLIOSTIRTOFEDCTVLTIAHRLNTIMD 1487
Qy      1494  YTRVIVLKGELQEGCAPSDLLQORGLFYSMAKDSGL 1530
Db      1488  YTRVIVLDRGEVCECDSPDNLQAKGLFYSMAKDSGL 1524

```

Search completed: December 15, 2005, 15:33:58
 Job time : 260.777 secs